

CONVERGENCE OF THE CASCADE ALGORITHM AT IRREGULAR SCALING FUNCTIONS

OLA BRATTELI AND PALLE E. T. JORGENSEN

ABSTRACT. The spectral properties of the Ruelle transfer operator which arises from a given polynomial wavelet filter are related to the convergence question for the cascade algorithm for approximation of the corresponding wavelet scaling function.

CONTENTS

| | |
|---------------------------------|----|
| 1. Introduction | 1 |
| 2. General Theory | 2 |
| 3. Some examples | 14 |
| List of Figures | 20 |
| 4. Conclusions | 33 |
| Appendix (by Brian F. Treadway) | 34 |
| References | 37 |

1. INTRODUCTION

Two operators from wavelet theory are studied: the refinement operator (alias the cascade approximation operator) M (see (2.9)), and the transfer operator R (see (2.25)). In the case of a compactly supported scaling function φ , we then consider the approximation problem

$$(1.1) \quad \lim_{n \rightarrow \infty} M^n \psi^{(0)} = \varphi \quad \text{in } L^2(\mathbb{R}),$$

where $\psi^{(0)} \in L^2(\mathbb{R})$ is given. A result is proved which relates the spectrum of R to the question of when $\varphi = \lim_{n \rightarrow \infty} M^n \psi^{(0)}$. For those vectors $\psi^{(0)}$ where the approximation holds, the rapidity of the approximation is related to the spectral data for R .

1991 *Mathematics Subject Classification*. Primary 46L60, 47D25, 42A16, 43A65; Secondary 46L45, 42A65, 41A15.

Key words and phrases. Wavelet, cascade algorithm, refinement operator, representation, orthogonal expansion, quadrature mirror filter, isometry in Hilbert space.

Research supported by the University of Oslo and the U.S. National Science Foundation.

2. GENERAL THEORY

It is well known that compactly supported scaling functions φ of a multiresolution analysis satisfy the functional equation

$$(2.1) \quad \varphi(x) = \sqrt{2} \sum_{k=0}^N a_k \varphi(2x - k)$$

(at least after an integer translation of φ) [Dau92, CoRy95]. The standard requirement that $\{\varphi(\cdot - k) \mid k \in \mathbb{Z}\}$ forms an orthonormal set of functions in $L^2(\mathbb{R})$ implies the conditions

$$(2.2) \quad \sum_{k \in \mathbb{Z}} \bar{a}_k a_{k+2l} = \delta_l$$

for all $l \in \mathbb{Z}$, and the second standard requirement that $\hat{\varphi}(0) = 1$ translates into the condition

$$(2.3) \quad \sum_{k \in \mathbb{Z}} a_k = \sqrt{2}.$$

If

$$(2.4) \quad m_0(z) = \sum_{k \in \mathbb{Z}} a_k z^k$$

for $z = e^{-it} \in \mathbb{T}$, condition (2.2) is equivalent to

$$(2.5) \quad |m_0(z)|^2 + |m_0(-z)|^2 = 2,$$

and (2.3) is equivalent to

$$(2.6) \quad m_0(1) = \sqrt{2}.$$

However, in general, orthogonality of $\{\varphi(x - k) \mid k \in \mathbb{Z}\}$ is a condition which is more restrictive than either one of the two equivalent conditions (2.2) or (2.5). The Fourier transform of (2.1) is

$$(2.7) \quad \hat{\varphi}(t) = \frac{1}{\sqrt{2}} m_0\left(\frac{t}{2}\right) \hat{\varphi}\left(\frac{t}{2}\right).$$

We apologize to engineers for using t to denote frequency. Since φ has compact support (and then (2.1) implies that the support is in $[0, N]$), $\hat{\varphi}$ is continuous at 0 and an iteration of (2.7) gives

$$(2.8) \quad \hat{\varphi}(t) = \prod_{k=1}^{\infty} \left(\frac{m_0(t2^{-k})}{\sqrt{2}} \right).$$

Since m_0 is a polynomial, this expansion converges uniformly on compacts.

Now, let $\psi^{(0)}$ be any bounded function of compact support such that $\widehat{\psi^{(0)}}(0) = 1$, and define by iteration

$$(2.9) \quad \begin{aligned} \psi^{(n+1)}(x) &= \left(M\psi^{(n)} \right)(x) \\ &= \sqrt{2} \sum_{k=0}^N a_k \psi^{(n)}(2x - k). \end{aligned}$$

Then

$$(2.10) \quad \widehat{\psi^{(n+1)}}(t) = \frac{1}{\sqrt{2}} m_0 \left(\frac{t}{2} \right) \widehat{\psi^{(n)}} \left(\frac{t}{2} \right),$$

and hence

$$(2.11) \quad \widehat{\psi^{(n)}}(t) = \prod_{k=1}^n \left(\frac{m_0(t2^{-k})}{\sqrt{2}} \right) \widehat{\psi^{(0)}}(t2^{-n}).$$

As a result, $\widehat{\psi^{(n)}} \xrightarrow[n \rightarrow \infty]{} \hat{\varphi}$, uniformly on compacts, and thus $\psi^{(n)} \xrightarrow[n \rightarrow \infty]{} \varphi$ in the distribution sense. In short, if $\{a_k \mid k = 0, \dots, N\}$ is a finite set of coefficients satisfying (2.3) alone, the refinement equation (2.1) possesses a distribution solution φ with $\hat{\varphi}(0) = 1$ and compact support in $[0, N]$. This solution is defined by (2.8), and is given as the distribution limit

$$(2.12) \quad \varphi = \lim_{n \rightarrow \infty} M^n \psi^{(0)},$$

where $\psi^{(0)}$ is any integrable function of compact support such that $\widehat{\psi^{(0)}}(0) = 1$. When this is used to depict the graph of φ , it is common to take $\psi^{(0)}$ to be the Haar function

$$(2.13) \quad \psi^{(0)}(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq 1, \\ 0 & \text{otherwise;} \end{cases}$$

see [Coh92]. In this paper, we discuss a variety of choices. The question of when the convergence in (2.12) is stronger than distribution convergence, and the related question of what regularity properties the limit φ has, have received much attention in the literature; see for example [Str96], [Vil94], [CoDa96], [Pol92], and references cited therein. In all these references, it is also assumed that the condition (2.2), alias (2.5), holds. If

$$(2.14) \quad \hat{\varphi}_n(t) = \prod_{k=1}^n \left(\frac{m_0(t2^{-k})}{\sqrt{2}} \right)$$

is the n 'th partial product in (2.8) one then shows, using (2.5), that

$$(2.15) \quad \begin{aligned} \int_{-2^n \pi}^{2^n \pi} |\hat{\varphi}_n(t)|^2 dt &= \int_0^{2^{n+1} \pi} |\hat{\varphi}_n(t)|^2 dt \\ &= \int_0^{2^n \pi} |\hat{\varphi}_{n-1}(t)|^2 \frac{1}{2} \left(|m_0(2^{-n}t)|^2 + |m_0(2^{-n}t + \pi)|^2 \right) dt \\ &= \int_0^{2^n \pi} |\hat{\varphi}_{n-1}(t)|^2 dt = \dots = \int_{-\pi}^{\pi} |\hat{\varphi}_0(t)|^2 dt = 2\pi. \end{aligned}$$

Since the convergence $\hat{\varphi}_n \xrightarrow[n \rightarrow \infty]{} \hat{\varphi}$ is uniform on compacts, it follows that

$$(2.16) \quad \|\hat{\varphi}\|_2^2 \leq 2\pi.$$

Hence $\varphi \in L^2(\mathbb{R})$, and $\|\varphi\|_2 \leq 1$. Since $\hat{\varphi}_n(\cdot) \chi_{[-\pi, \pi]}(\cdot 2^{-n})$ converges, as $n \rightarrow \infty$, uniformly on compacts, and has constant L^2 -norm equal to $\sqrt{2\pi}$ by (2.15), it follows

that this sequence converges weakly to $\hat{\varphi}$ in $L^2(\mathbb{R})$. Consequently it converges in L^2 -norm to $\hat{\varphi}$ if and only if

$$(2.17) \quad \|\hat{\varphi}\|_2^2 = 2\pi,$$

i.e., if and only if

$$(2.18) \quad \|\varphi\|_2 = 1.$$

It is well known that this is equivalent to several other more practical conditions, like (2.19)–(2.23) below:

$$(2.19) \quad \{\varphi(\cdot - k)\}_{k \in \mathbb{Z}} \text{ is an orthonormal set in } L^2(\mathbb{R}).$$

$$(2.20) \quad \sum_{k \in \mathbb{Z}} |\hat{\varphi}(t + 2\pi k)|^2 = 1 \text{ for all } t \in \mathbb{R}.$$

$$(2.21) \quad \text{The only trigonometric polynomials } \xi \text{ satisfying}$$

$$\xi(z) = \frac{1}{2} \sum_{w^2=z} |m_0(w)|^2 \xi(w)$$

are the constants.

$$(2.22) \quad \text{There is no nontrivial cycle in}$$

$$\{z \in \mathbb{T} \mid |m_0(z)| = \sqrt{2}\} = \{z \in \mathbb{T} \mid |m_0(-z)| = 0\}$$

for the doubling map $z \mapsto z^2$.

See for example [Hor95, Theorem 3.3.6] or [CoRy95, Chapter 2] for details. We will especially need (2.21) in the sequel. By the previous remarks we also note that this is equivalent to:

$$(2.23) \quad \text{The cascade algorithm, with}$$

$$\psi^{(0)}(x) = \check{\chi}_{[-\pi, \pi]}(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{itx} dt = \frac{1}{\pi x} \sin(\pi x),$$

converges in L^2 -norm to φ ;

and then $\{\varphi(\cdot - k)\}$ is an orthonormal set.

In contrast to the distribution convergence, this latter convergence depends very sensitively on the choice of initial function $\psi^{(0)}$. If for example the $\psi^{(0)}$ starting vector above is replaced by $\frac{1}{N\pi} \sin(N\pi x)$, then $\|\psi^{(n)}\|_2^2 = N$ for all n , so $\psi^{(n)}$ cannot converge to φ in norm (although it does so weakly). The question is then which initial functions $\psi^{(0)}$ can be used. One approach, developed by Strang [Str96], establishes L^2 -convergence under general circumstances if $\psi^{(0)}$ is chosen such that

$$(2.24) \quad \left\{ \psi^{(0)}(\cdot - k) \right\}_{k \in \mathbb{Z}}$$

is an orthonormal set. Thus the standard choice $\psi^{(0)} = \chi_{[0,1]}$ is included. To describe these “general circumstances” we introduce the Ruelle operator

$$(2.25) \quad (R\xi)(z) = \frac{1}{2} \sum_{w^2=z} |m_0(w)|^2 \xi(w).$$

We may view R as an operator on any of the spaces

$$(2.26) \quad \mathbb{C}[z, z^{-1}] \subset C(\mathbb{T}) \subset L^\infty(\mathbb{T}) \subset L^2(\mathbb{T}).$$

and in particular it is clear from (2.25), (2.4) that R maps any of these spaces into itself. Since $R(\mathbb{C}[z, z^{-1}]) \subset \mathbb{C}[z, z^{-1}]$, the invariance of $C(\mathbb{T})$ follows from

the Stone–Weierstrass theorem. This also follows directly from (2.25), using the continuity of m_0 .

If $P[n, m]$, $n \leq m$, is the subspace of $\mathbb{C}[z, z^{-1}]$ consisting of trigonometric polynomials of the form $\sum_{k=n}^m b_k z^k$, we note that

$$(2.27) \quad R(P[n, m]) \subset P\left[-\left(\frac{N-n}{2}\right), \left(\frac{m+N}{2}\right)\right]$$

where $[x]$ is the largest integer $\leq x$. Thus any $P[n, m]$ will ultimately be mapped into $P[-N, N]$ by repeated applications of R , so all the spaces

$$(2.28) \quad P[-N, N] \subset \mathbb{C}[z, z^{-1}] \subset C(\mathbb{T}) \subset L^\infty(\mathbb{T}) \subset L^2(\mathbb{T})$$

are invariant under R .

Let us study the spectral properties and the norm of R on the various spaces. Note that $\xi(z) = 1$ (the constant function $\mathbb{1}$) is an eigenvector of R with eigenvalue 1 in all the subspaces; and, by (2.21), it is the unique eigenvector up to a scalar with eigenvalue 1 in $\mathbb{C}[z, z^{-1}]$ if and only if $\|\varphi\|_2 = 1$, i.e., if and only if the cascade (2.23) converges in L^2 -norm. In any case, (2.25) and (2.5) immediately imply that

$$(2.29) \quad \|R\|_{\infty \rightarrow \infty} = 1,$$

and hence the spectral radius of R , as an operator on the four left-hand subspaces in (2.28), is

$$(2.30) \quad \rho_\infty(R) = 1.$$

Let us note *en passant* that the behaviour of R as an operator on $L^2(\mathbb{T})$ is different. From (2.25), we get for the $L^2(\mathbb{T})$ -adjoint operator:

$$(2.31) \quad (R^* \xi)(z) = |m_0(z)|^2 \xi(z^2),$$

and hence

$$(2.32) \quad \begin{aligned} (RR^* \xi)(z) &= \frac{1}{2} \sum_{w^2=z} |m_0(w)|^2 (R^* \xi)(w) \\ &= \frac{1}{2} \sum_{w^2=z} |m_0(w)|^4 \xi(z). \end{aligned}$$

Thus RR^* is a multiplication operator, and it follows from (2.5) and (2.6) that

$$(2.33) \quad \|R\| = \|RR^*\|^{\frac{1}{2}} = \left(\frac{1}{2} (\sqrt{2})^4\right)^{\frac{1}{2}} = \sqrt{2},$$

so R is not contractive, nor normal, on $L^2(\mathbb{T})$.

What is the spectral radius? From (2.25) and (2.31), we verify the identities

$$(2.34) \quad (R^n \xi)(z) = \frac{1}{2^n} \sum_{w^{2^n}=z} \prod_{k=0}^{n-1} |m_0(w^{2^k})|^2 \xi(w)$$

and

$$(2.35) \quad (R^{*n} \xi)(z) = \prod_{k=0}^{n-1} |m_0(z^{2^k})|^2 \xi(z^{2^n}),$$

and hence

$$(2.36) \quad (R^n R^{*n} \xi)(z) = \frac{1}{2^n} \sum_{w^{2^n}=z} \prod_{k=0}^{n-1} |m_0(w^{2^k})|^4 \xi(z)$$

holds. Thus $R^n R^{*n}$ is the operator of multiplication by the function

$$(2.37) \quad \frac{1}{2^n} \sum_{w^{2^n}=z} \prod_{k=0}^{n-1} |m_0(w^{2^k})|^4 = p_n(z),$$

and it follows that the spectral radius of R is

$$(2.38) \quad \rho_2(R) = \lim_{n \rightarrow \infty} \|p_n\|_{\infty}^{\frac{1}{2^n}}.$$

Note that

$$(2.39) \quad p_n(z) = \frac{1}{2} \sum_{w^2=z} |m_0(w)|^4 p_{n-1}(w)$$

and

$$(2.40) \quad p_0(z) = 1.$$

But, by (2.5) and (2.6), we have

$$(2.41) \quad 1 \leq \frac{1}{2} \sum_{w^2=z} |m_0(w)|^4 \leq 2,$$

with both extremities attained. By (2.39) and (2.41), and induction, we further have

$$(2.42) \quad 1 \leq p_n(z) \leq 2^n,$$

and hence, by (2.38),

$$(2.43) \quad 1 \leq \rho_2(R) \leq \sqrt{2}.$$

After this diversion into Hilbert space, let us return to the cascade algorithm. Strang's criterion is based on the following basic lemma: If ψ_1, ψ_2 are L^2 -functions of bounded support, define their relative polynomial by

$$(2.44) \quad p(\psi_1, \psi_2)(z) = \sum_k z^k \int_{\mathbb{R}} \overline{\psi_1(x-k)} \psi_2(x) dx.$$

Note that

$$p(L^2([0, N]), L^2([0, N])) \subset P[-N, N],$$

i.e., p maps $L^2([0, N]) \times L^2([0, N])$ into a $2N+1$ -dimensional space of trigonometric polynomials.

Lemma 2.1. *If M is the cascade transform defined by (2.9), then*

$$(2.45) \quad p(M\psi_1, M\psi_2) = R(p(\psi_1, \psi_2))$$

for any pair ψ_1, ψ_2 of L^2 -functions of compact support, where R is the Ruelle operator (2.25) on $\mathbb{C}[z, z^{-1}]$.

Proof. If $\xi(z) = \sum_k x_k z^k \in \mathbb{C}[z, z^{-1}]$, its Fourier transform is $\langle x_k \rangle_{k \in \mathbb{Z}}$, and the Ruelle operator (2.25) transforms into

$$(2.46) \quad (\hat{R}x)_k = \sum_j \sum_l \bar{a}_j a_l x_{j-l+2k}.$$

This follows from the next computation on

$$(2.47) \quad \xi(z) = \sum_k x_k z^k.$$

We have

$$(2.48) \quad \begin{aligned} (R\xi)(z) &= \frac{1}{2} \sum_{w^2=z} \overline{m_0(w)} m_0(w) \xi(w) \\ &= \frac{1}{2} \sum_{w^2=z} \sum_{jl} \bar{a}_j a_l w^{-j} w^l \sum_k x_k w^k \\ &= \frac{1}{2} \sum_{w^2=z} \sum_{jlk} \bar{a}_j a_l x_k w^{-j+l+k} \\ &= \frac{1}{2} \sum_{w^2=z} \sum_{jln} \bar{a}_j a_l x_{j-l+n} w^n \\ &= \frac{1}{2} \sum_n \sum_{w^2=z} w^n \sum_{jl} \bar{a}_j a_l x_{j-l+n}. \end{aligned}$$

The terms with odd n disappear in the last sum, so

$$(2.49) \quad (R\xi)(z) = \sum_k z^k \sum_{jl} \bar{a}_j a_l x_{j-l+2k},$$

and (2.46) follows.

If ψ_1, ψ_2 are L^2 -functions with compact support, define

$$(2.50) \quad b_k = \int_{\mathbb{R}} \overline{\psi_1(x-k)} \psi_2(x) dx,$$

$$(2.51) \quad c_k = \int_{\mathbb{R}} \overline{(M\psi_1)(x-k)} (M\psi_2)(x) dx.$$

To prove (2.45), we have to show

$$(2.52) \quad c_k = (\hat{R}b)_k.$$

We have

$$\begin{aligned}
 (2.53) \quad c_k &= \int_{\mathbb{R}} \overline{(M\psi_1)(x-k)} (M\psi_2)(x) dx \\
 &= 2 \int_{\mathbb{R}} \sum_{mn} \overline{a_m \psi_1(2x-2k-m)} a_n \psi_2(2x-n) dx \\
 &= \int_{\mathbb{R}} \sum_{mn} \bar{a}_m a_n \overline{\psi_1(x-2k-m)} \psi_2(x-n) dx \\
 &= \int_{\mathbb{R}} \sum_{mn} \bar{a}_m a_n \overline{\psi_1(x+n-m-2k)} \psi_2(x) dx \\
 &= \sum_{mn} \bar{a}_m a_n b_{m-n+2k} = \left(\hat{R}b \right)_k.
 \end{aligned}$$

This shows (2.52), and Lemma 2.1 is proved. \square

We will list a few more preliminaries on the relative polynomials before we prove our theorem on spectrum and cascade approximation. As in Lemma 2.1, we consider functions ψ, ψ_1, ψ_2 in $L^2(\mathbb{R})$ of compact support. Usually the support of the functions in question will be assumed contained in a fixed interval $[0, N] \subset \mathbb{R}$. Consider $\xi(z) = \sum_k \xi_k z^k \in \mathbb{C}[z, z^{-1}] \subset C(\mathbb{T})$, $\xi_k = \hat{\xi}(k)$ denoting the Fourier coefficients. Defining

$$(2.54) \quad (\xi * \psi)(x) := \sum_{k \in \mathbb{Z}} \xi_k \psi(x-k),$$

we note that

$$(2.55) \quad (\xi * \psi)^{\wedge}(t) = \xi(e^{-it}) \hat{\psi}(t),$$

where $\hat{\cdot}$ as usual refers to the \mathbb{R} -Fourier transform. The sesquilinear operator $p(\psi_1, \psi_2)$ in (2.44) will be viewed as a quadratic form

$$(2.56) \quad p: L^2(\mathbb{R})_c \times L^2(\mathbb{R})_c \longrightarrow \mathbb{C}[z, z^{-1}] \subset C(\mathbb{T}) \subset L^2(\mathbb{T}),$$

with the subscript c standing for compact support.

Positivity of this $C(\mathbb{T})$ -valued form p follows from the following identity:

$$\begin{aligned}
 (2.57) \quad p(\psi, \psi)(z) &= \sum_{k \in \mathbb{Z}} z^k \int_{\mathbb{R}} \overline{\psi(x-k)} \psi(x) dx \\
 &= \sum_k z^k \frac{1}{2\pi} \int_0^{2\pi} e_k(t) \text{PER} \left(\left| \hat{\psi} \right|^2 \right) (e^{it}) dt \\
 &= \text{PER} \left(\left| \hat{\psi} \right|^2 \right) (z),
 \end{aligned}$$

with the convention $z = e^{-it}$, $e_k(t) = z^k = e^{-ikt}$, and

$$\text{PER}(f)(e^{it}) = \sum_{n \in \mathbb{Z}} f(t + 2\pi n).$$

Similarly we derive the formula

$$(2.58) \quad p(\psi_1, \psi_2)(z) = \text{PER} \left(\hat{\psi}_1 \hat{\psi}_2 \right) (z).$$

From this it is immediate that the relative polynomial p has the sesquilinearity property

$$(2.59) \quad p(\psi_1, \psi_2)(z) = \overline{p(\psi_2, \psi_1)(z)},$$

and that it is densely defined on $L^2(\mathbb{R}) \times L^2(\mathbb{R})$. We next show that when $\psi_1 \in L^2(\mathbb{R})_c$ is given, then $p(\psi_1, \cdot)$ extends boundedly to $L^2(\mathbb{R})$ in the second variable, and that it becomes a module mapping in that variable relative to $\mathbb{C}[z, z^{-1}] \subset L^\infty(\mathbb{T})$. This is actually slightly more than what we need (see the proof of (ii) \Rightarrow (i) in Theorem 2.5). On the other hand, one may use the Zak transform [Dau92] to prove a stronger continuity estimate: p may be extended to a sesquilinear function $L^2(\mathbb{R}) \times L^2(\mathbb{R}) \rightarrow L^1(\mathbb{T})$ such that

$$\|p(\psi_1, \psi_2)\|_1 \leq \|\psi_1\|_2 \cdot \|\psi_2\|_2.$$

We turn to the details.

Lemma 2.2. *The relative polynomial p has the following properties:*

(a) *Norm estimate:*

$$(2.60) \quad \begin{aligned} \|p(\psi_1, \psi_2)\|_{L^2(\mathbb{T})} &= \frac{1}{2\pi} \int_0^{2\pi} \left| \text{PER}(\hat{\psi}_1 \hat{\psi}_2)(t) \right|^2 dt \\ &\leq \left\| \text{PER} |\hat{\psi}_1|^2 \right\|_\infty \cdot \|\psi_2\|_{L^2(\mathbb{R})}^2. \end{aligned}$$

(b) *Module property:*

$$(2.61) \quad p(\psi_1, \xi * \psi_2)(z) = \xi(z) p(\psi_1, \psi_2)(z).$$

(c) *Ruelle operator covariance:*

$$(2.62) \quad M_{m_0}(\xi * \psi) = M_{m_0 \xi}(\psi)$$

and thus

$$(2.63) \quad R(p(\psi_1, \xi * \psi_2))(z) = p(M_{m_0} \psi_1, M_{\xi m_0} \psi_2)(z),$$

where ξm_0 is the pointwise product.

(d) *Cascade smoothing:*

$$(2.64) \quad (M^n(\xi * \psi))^\wedge(t) = \xi\left(e^{-\frac{it}{2^n}}\right) (\widehat{M^n \psi})(t),$$

where $M = M_{m_0}$ is the cascade operator.

Proof. *Ad (a):* By Parseval's identity in $L^2(\mathbb{T})$,

$$\begin{aligned}
 (2.65) \quad \|p(\psi_1, \psi_2)\|_{L^2(\mathbb{T})}^2 &= \frac{1}{2\pi} \int_0^{2\pi} |p(\psi_1, \psi_2)(e^{-it})|^2 dt \\
 &= \sum_{k \in \mathbb{Z}} \left| \int_{\mathbb{R}} \overline{\psi_1(x-k)} \psi_2(x) dx \right|^2 \\
 &= \sum_{k \in \mathbb{Z}} \left| \frac{1}{2\pi} \int_{\mathbb{R}} e_k(t) \bar{\hat{\psi}}_1(t) \hat{\psi}_2(t) dt \right|^2 \\
 &= \sum_{k \in \mathbb{Z}} \left| \frac{1}{2\pi} \int_0^{2\pi} e_k(t) \text{PER}(\bar{\hat{\psi}}_1 \hat{\psi}_2)(t) dt \right|^2 \\
 &= \frac{1}{2\pi} \int_0^{2\pi} |\text{PER}(\bar{\hat{\psi}}_1 \hat{\psi}_2)|^2 dt \\
 &\leq \frac{1}{2\pi} \int_0^{2\pi} \left(\text{PER}|\hat{\psi}_1|^2 \cdot \text{PER}|\hat{\psi}_2|^2 \right) dt \\
 &\leq \left\| \text{PER}|\hat{\psi}_1|^2 \right\|_{\infty} \cdot \frac{1}{2\pi} \int_0^{2\pi} \text{PER}|\hat{\psi}_2|^2 dt \\
 &= \left\| \text{PER}|\hat{\psi}_1|^2 \right\|_{\infty} \cdot \frac{1}{2\pi} \int_{\mathbb{R}} |\hat{\psi}_2|^2 dt \\
 &= \left\| \text{PER}|\hat{\psi}_1|^2 \right\|_{\infty} \cdot \int_{\mathbb{R}} |\psi_2(x)|^2 dx.
 \end{aligned}$$

Ad (b): We have

$$\begin{aligned}
 (2.66) \quad p(\psi_1, \xi * \psi_2)(z) &= \sum_{k \in \mathbb{Z}} z^k \int_{\mathbb{R}} \overline{\psi_1(x-k)} (\xi * \psi_2)(x) dx \\
 &= \sum_k \sum_l \xi_l z^k \int_{\mathbb{R}} \overline{\psi_1(x-k)} \psi_2(x-l) dx \\
 &= \sum_k \sum_l \xi_l z^l z^{k-l} \int_{\mathbb{R}} \overline{\psi_1(x+l-k)} \psi_2(x) dx \\
 &= \xi(z) p(\psi_1, \psi_2)(z).
 \end{aligned}$$

Ad (c): Since

$$(2.67) \quad R(p(\psi_1, \xi * \psi_2)) = p(M\psi_1, M(\xi * \psi_2))$$

by Lemma 2.1, we need only calculate $M(\xi * \psi_2)$, where M is the cascade operator $M = M_{m_0}$ given by the low-pass filter m_0 .

$$\begin{aligned}
 (2.68) \quad M(\xi * \psi_2)(x) &= \sqrt{2} \sum_k a_k (\xi * \psi_2)(2x - k) \\
 &= \sqrt{2} \sum_k a_k \sum_l \xi_l \psi_2(2x - k - l) \\
 &= \sqrt{2} \sum_n \sum_k a_k \xi_{n-k} \psi_2(2x - n) \\
 &= \sum_n (m_0 \xi)^\sim(n) \psi_2(2x - n) \\
 &= M_{m_0 \xi} \psi_2(x),
 \end{aligned}$$

where $M_{m_0 \xi}$ is the cascade operator corresponding to the product filter $(m_0 \xi)(z) = m_0(z) \xi(z)$.

Ad (d): From the definition of the cascade operator M , we get $(M\psi)^\sim(t) = \frac{1}{\sqrt{2}} m_0\left(\frac{t}{2}\right) \hat{\psi}\left(\frac{t}{2}\right)$. Now apply this to $\xi * \psi$, and use $(\xi * \psi)^\sim(t) = \xi(e^{-it}) \hat{\psi}(t)$. Iteration yields

$$\begin{aligned}
 (2.69) \quad (M^n(\xi * \psi))^\sim(t) &= \prod_{k=1}^n \frac{1}{\sqrt{2}} m_0\left(\frac{t}{2^k}\right) (\xi * \psi)^\sim\left(\frac{t}{2^n}\right) \\
 &= \xi\left(e^{-\frac{it}{2^n}}\right) \prod_{k=1}^n \frac{1}{\sqrt{2}} m_0\left(\frac{t}{2^k}\right) \hat{\psi}\left(\frac{t}{2^n}\right) \\
 &= \xi\left(e^{-\frac{it}{2^n}}\right) (M^n \psi)^\sim(t). \quad \square
 \end{aligned}$$

As a corollary, we note that

Corollary 2.3.

$$(2.70) \quad p(L^2(\mathbb{R})_c \times L^2(\mathbb{R})_c) = \mathbb{C}[z, z^{-1}].$$

Proof. We have already commented on one inclusion in (2.56), and the second follows from

$$\xi = \xi \cdot p\left(\psi^{(0)}, \psi^{(0)}\right) = p\left(\psi^{(0)}, \xi * \psi^{(0)}\right),$$

where $\psi^{(0)} \in L^2(\mathbb{R})_c$ is chosen such that $p(\psi^{(0)}, \psi^{(0)}) = \mathbf{1}$.

$$(\text{Equivalently, } \text{PER}\left(\overline{\psi^{(0)}} \psi^{(0)}\right) = \text{PER}\left(\left|\widehat{\psi^{(0)}}\right|^2\right) \equiv 1.) \quad \square$$

Remark 2.4. The positive elements in $\mathbb{C}[z, z^{-1}]$ (i.e., pointwise nonnegative as functions on \mathbb{T}) are of the form $|\xi|^2 = \bar{\xi}\xi$ by a theorem of Riesz [Dau92, Chapter 6], [Akh65, p. 181], and we get $|\xi|^2 = p(\xi * \psi^{(0)}, \xi * \psi^{(0)})$.

We are now ready to prove the main theorem of convergence of the cascade algorithm, which is a version of Theorem 4 in [Str96].

Theorem 2.5. *Let a_0, a_1, \dots, a_N be complex numbers satisfying (2.2)–(2.3) and let φ be the associated scaling function defined by (2.8). Identify the Ruelle operator R given in (2.25) with its restriction to $P[-N, N]$ (or to any $P[n, m]$ with $n \leq -N$, $m \geq N$). The following conditions are equivalent.*

- (i) R has 1 as a simple eigenvalue and $|\lambda| < 1$ for all other eigenvalues λ of R .
(ii) If $\psi^{(0)} \in L^2(\mathbb{R})$ is a function with compact support such that $\{\psi^{(0)}(\cdot - k)\}_{k \in \mathbb{Z}}$ is an orthonormal set and $\widehat{\psi^{(0)}}(0) = 1$ then

$$(2.71) \quad \lim_{n \rightarrow \infty} \|\varphi - M^n \psi^{(0)}\|_2 = 0.$$

Proof. (i) \Rightarrow (ii). As an aside, remark that R having 1 as a simple eigenvalue means here that the corresponding eigenspace is one-dimensional. But since $\|R\| = 1$, and hence $n \mapsto \|R^n\|'$ is bounded in any equivalent norm $\|\cdot\|'$ on the linear operators on $P[-N, N] \cong \mathbb{C}^{2N+1}$, it follows from Jordan's theorem that the multiplicity of 1 in the characteristic polynomial is 1 too.

Let us view $P[-N, N] = \left\{ \sum_{k=-N}^N x_k z^k \right\}$ as the space of sequences $x = \langle x_k \rangle_{k=-N}^N$. Since

$$(2.72) \quad \begin{aligned} (R\xi)(1) &= \frac{1}{2} \left(|m_0(1)|^2 \xi(1) + |m_0(-1)|^2 \xi(-1) \right) \\ &= \xi(1) \end{aligned}$$

for all $\xi \in \mathbb{C}[z, z^{-1}]$, we have

$$(2.73) \quad \sum_k (\hat{R}x)(k) = \sum_k x(k).$$

But as $\hat{R}(\delta_0) = \delta_0$, δ_0 is the unique eigenvector of \hat{R} corresponding to eigenvalue 1, and as the functional $x \mapsto \sum_k x(k)$ is preserved by \hat{R} , it follows from (i) that

$$(2.74) \quad \lim_{n \rightarrow \infty} \hat{R}^n x = \left(\sum_k x(k) \right) \delta_0$$

for all finite sequences x . Thus, by Lemma 2.1,

$$(2.75) \quad \begin{aligned} p(\varphi, M^n \psi^{(0)}) &= p(M^n \varphi, M^n \psi^{(0)}) \\ &= R^n(p(\varphi, \psi^{(0)})) \\ &\longrightarrow \left(\sum_k \int_{\mathbb{R}} \overline{\varphi(x-k)} \psi^{(0)}(x) dx \right) \cdot \mathbf{1}. \end{aligned}$$

But the two assumptions on $\psi^{(0)}$ imply the so-called Strang-Fix condition [Vil94, Str96]

$$(2.76) \quad \sum_k \psi^{(0)}(x+k) = 1$$

for all x , and hence by the above,

$$(2.77) \quad p(\varphi, M^n \psi^{(0)}) \longrightarrow \int_{\mathbb{R}} \overline{\varphi(x)} dx = \overline{\hat{\varphi}(0)} = 1.$$

In particular this means that

$$(2.78) \quad \lim_{n \rightarrow \infty} \langle \varphi | M^n \psi^{(0)} \rangle = 1.$$

But by (2.45),

$$(2.79) \quad p(M^n \psi^{(0)}, M^n \psi^{(0)}) = R^n(p(\psi^{(0)}, \psi^{(0)})).$$

Since $p(\psi^{(0)}, \psi^{(0)}) = \mathbb{1}$ by orthonormality, we conclude that

$$(2.80) \quad p(M^n \psi^{(0)}, M^n \psi^{(0)}) = R^n \mathbb{1} = \mathbb{1},$$

and in particular,

$$(2.81) \quad \|M^n \psi^{(0)}\|_2^2 = 1$$

for all n . Also

$$(2.82) \quad \|M^n \varphi\|_2^2 = \|\varphi\|_2^2 = 1,$$

so, finally,

$$(2.83) \quad \begin{aligned} \|\varphi - M^n \psi^{(0)}\|_2^2 &= \|\varphi\|_2^2 - 2 \operatorname{Re} \langle \varphi | M^n \psi^{(0)} \rangle + \|\psi^{(0)}\|_2^2 \\ &\xrightarrow{n \rightarrow \infty} 1 - 2 + 1 = 0. \end{aligned}$$

(ii) \Rightarrow (i). We now assume cascade convergence in the sense (ii), i.e.,

$$(2.84) \quad \|\varphi - M^n \psi^{(0)}\|_2 \longrightarrow 0$$

for the initial vectors $\psi^{(0)}$ which are specified in (ii). The object is to derive from this the spectral picture for R as specified in (i), and R will be identified with its restriction to $P[-N, N]$ as mentioned. Of course $P[-N, N] \subset C(\mathbb{T})$, and R is also, by (2.28), an operator mapping $C(\mathbb{T})$ into itself. Its adjoint on the dual space of measures $M(\mathbb{T})$ is given by $(R^* \mu)(\xi) = \mu(R\xi) = \int_{\mathbb{T}} (R\xi)(z) d\mu(z)$, $\xi \in C(\mathbb{T})$. The Dirac point-measure $\delta_1 \in M(\mathbb{T})$, given by $\delta_1(\xi) = \xi(1)$, is invariant by (2.5), i.e., $R^*(\delta_1) = \delta_1$.

Consider the eigenvalue problem:

$$(2.85) \quad R\xi_0 = \lambda\xi_0, \quad \lambda \in \mathbb{C}, \quad \xi_0 \in P[-N, N] \setminus \{0\}.$$

Then $\delta_1(\xi_0) = \delta_1(R\xi_0) = \lambda\delta_1(\xi_0)$, so $\delta_1(\xi_0) = \xi_0(1) = 0$ if $\lambda \neq 1$. We assume this, and since $\|R\|_{\infty \rightarrow \infty} = 1$, the discussion may be restricted to $|\lambda| = 1$. We claim that, if $\lambda \neq 1$, $\lambda \in \mathbb{T}$, then $\xi_0 = 0$, so we cannot have nontrivial peripheral spectrum.

By (ii), $\|\varphi - M^n \psi^{(0)}\|_2 \rightarrow 0$, where $\psi^{(0)}$ is any initial vector with the stated conditions, e.g., $\psi^{(0)} = \chi_{[0,1]}$. Using Lemma 2.2(d), we also get

$$(2.86) \quad \left\| \varphi - M^n \left(\xi * \psi^{(0)} \right) \right\|_{L^2(\mathbb{R})} \longrightarrow 0,$$

whenever $\xi(1) = 1$. For example, take $\xi = \mathbb{1} + c\xi_0$ to have this satisfied. Then $R^n \xi = \mathbb{1} + c\lambda^n \xi_0$, i.e., a divergent sequence if $\lambda \neq 1$ and $c \neq 0$, supposing $\xi_0 \neq 0$.

Using Lemma 2.2, we will show that

$$(2.87) \quad \begin{aligned} R^n \xi &= p \left(M^n \psi^{(0)}, M^n \left(\xi * \psi^{(0)} \right) \right) \\ &\xrightarrow{n \rightarrow \infty} p(\varphi, \varphi), \end{aligned}$$

where the last convergence is in the finite-dimensional subspace of $\mathbb{C}[z, z^{-1}]$ and thus in any norm. This will contradict the divergence of $R^n \xi$. The formula (2.87) can be verified in two ways: since φ , $M^n(\psi^{(0)})$ and $M^n(\xi * \psi^{(0)})$ all have support inside a common compact set, the convergence is immediate from the finite sum (2.44), (2.84), and (2.86). Alternatively one can use Lemma 2.2(a),

and $\text{PER} \left| \widehat{\psi^{(0)}} \right|^2 = \mathbf{1}$. In checking the conditions in Lemma 2.2(a), we note that $M^n (\xi * \psi^{(0)}) \xrightarrow{n \rightarrow \infty} \xi(1) \varphi$, in $L^2(\mathbb{R})$, so we must verify that

$$\left\| \text{PER} \left(\left| \widehat{M^n \psi^{(0)}} \right|^2 \right) \right\|_{\infty}$$

is bounded in n . But the function inside $\|\cdot\|_{\infty}$ equals

$$p \left(M^n \psi^{(0)}, M^n \psi^{(0)} \right) = R^n \left(p \left(\psi^{(0)}, \psi^{(0)} \right) \right) = R^n \mathbf{1} = \mathbf{1}.$$

This contradiction completes the first part of the proof of (ii) \Rightarrow (i).

It remains to show that (ii) of Theorem 2.5 implies that $\lambda = 1$ has multiplicity one in the spectrum of R , where again R is identified with its restriction to $P[-N, N]$. Since $R\mathbf{1} = \mathbf{1}$, we need only exclude that the multiplicity is 2 or more. But $\|M^n \psi^{(0)}\|_2 = 1$ for all n by (2.45) and $R\mathbf{1} = \mathbf{1}$ and it follows from (2.71) that $\|\varphi\|_2 = 1$. By (2.18) \Rightarrow (2.21) it follows that $\lambda = 1$ has multiplicity 1. \square

3. SOME EXAMPLES

Our interest in the subject of cascade approximation was ignited when using the cascade algorithm in [BEJ99] to draw the scaling function associated with the low-pass wavelet filter

$$(3.1) \quad m_0^{(\theta)}(z) = \sum_{k=0}^3 a_k^{(\theta)} z^k$$

where

$$(3.2) \quad \begin{aligned} a_0^{(\theta)} &= \frac{1}{2\sqrt{2}} (1 - \cos \theta + \sin \theta), \\ a_1^{(\theta)} &= \frac{1}{2\sqrt{2}} (1 - \cos \theta - \sin \theta), \\ a_2^{(\theta)} &= \frac{1}{2\sqrt{2}} (1 + \cos \theta - \sin \theta), \\ a_3^{(\theta)} &= \frac{1}{2\sqrt{2}} (1 + \cos \theta + \sin \theta); \end{aligned}$$

and θ varies over the circle; see [Pol89, Pol90]. This family is discussed in detail in [BEJ99]. The symmetry $\varphi^{(\pi-\theta)}(x) = \varphi^{(\theta)}(3-x)$ (see [BEJ99, Proposition 4.1]) implies that we may limit the discussion to the interval $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$. In this interval the conditions (2.17)–(2.22) are fulfilled with the one exception of $\theta = \frac{\pi}{2}$, where

$$m_0^{(\frac{\pi}{2})}(z) = \frac{1}{\sqrt{2}} (1 + z^3)$$

and the scaling function φ is given by

$$\varphi^{(\frac{\pi}{2})}(x) = \begin{cases} \frac{1}{3} & \text{for } 0 \leq x \leq 3, \\ 0 & \text{otherwise.} \end{cases}$$

It has been observed by several authors that at this point the cascade approximants converge merely weakly; see [BEJ99, Figure 6], [CoRy95, Figure 3.3], [Str96, Note 4], [Coh92, Figure 4]. When using the cascade algorithm to depict $\varphi^{(\theta)}$ for θ near $\frac{\pi}{2}$, this weak convergence seems to persist; see Figures 5–6. So let us compute

the Ruelle operator in (2.25) R as a 7×7 matrix on $P[-3, 3]$. The result is the following slant-Toeplitz matrix:

$$R = \begin{bmatrix} b & 0 & 0 & 0 & 0 & 0 & 0 \\ c & 0 & b & 0 & 0 & 0 & 0 \\ c & 1 & c & 0 & b & 0 & 0 \\ b & 0 & c & 1 & c & 0 & b \\ 0 & 0 & b & 0 & c & 1 & c \\ 0 & 0 & 0 & 0 & b & 0 & c \\ 0 & 0 & 0 & 0 & 0 & 0 & b \end{bmatrix}$$

where $b = a_3 a_0$ and $c = a_1 a_0 + a_2 a_1 + a_3 a_2$, and we have used (2.2). Viewing $a = (a_0, a_1, a_2, a_3)$ as a function on \mathbb{Z}_4 , and letting T denote cyclic translation on \mathbb{Z}_4 ,

$$T(a_0, a_1, a_2, a_3) = (a_1, a_2, a_3, a_0),$$

the relations (2.2) take the form

$$\langle a, a \rangle = 1, \quad \langle a, T^2 a \rangle = 0.$$

Furthermore

$$\langle a, Ta \rangle = \langle a, T^3 a \rangle = b + c,$$

and by (2.3),

$$\sum_{k=0}^3 T^k a = \left(\sum a_i, \sum a_i, \sum a_i, \sum a_i \right) = \sqrt{2}(1, 1, 1, 1).$$

Thus

$$\left\langle a, \sum_{k=0}^3 T^k a \right\rangle = \left\langle (a_0, a_1, a_2, a_3), \sqrt{2}(1, 1, 1, 1) \right\rangle = \sqrt{2} \sum_i a_i = 2.$$

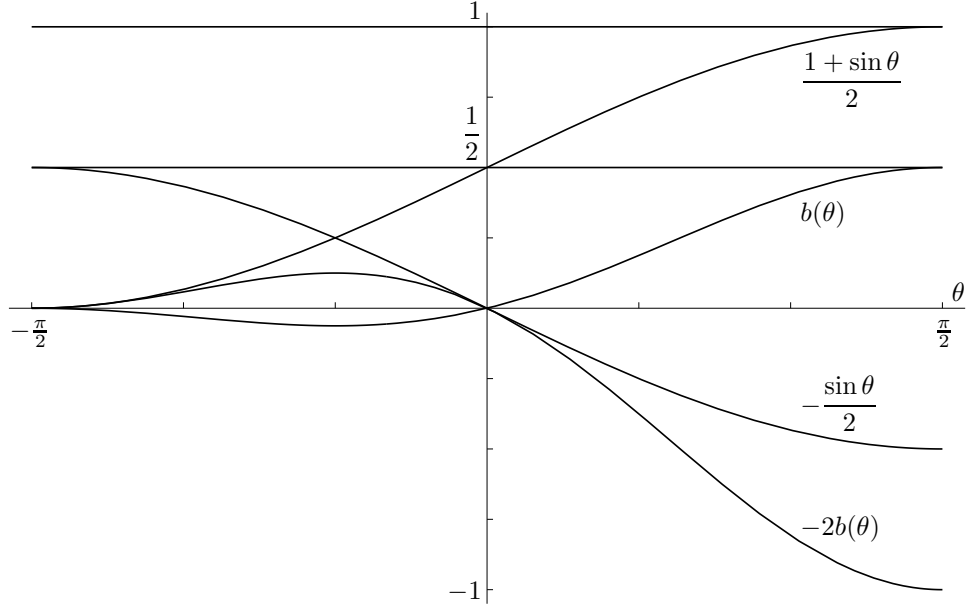
On the other side,

$$\left\langle a, \sum_{k=0}^3 T^k a \right\rangle = \sum_{k=0}^3 \langle a, T^k a \rangle = 1 + (b + c) + 0 + (b + c) = 1 + 2(b + c),$$

so

$$b + c = 1/2.$$

Since $b(\theta) = \frac{1}{8}[1 + 2\sin\theta - \cos(2\theta)]$, $b(\theta)$ ranges over $[-\frac{1}{16}, \frac{1}{2}]$ as θ ranges over $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The characteristic equation of R has the obvious roots 1 and b , where b has multiplicity 2. Using the flip symmetry of the matrix under the joint flip around the fourth column and fourth row, the remaining fourth-order polynomial factors into two second-order polynomials, and thus one deduces that the eigenvalues of R

FIGURE 1. Eigenvalues $\lambda(\theta)$ of R

are

| Eigenvalue $\lambda(\theta)$ | Range of $\lambda(\theta)$ |
|---|--------------------------------|
| 1 | $\{1\}$ |
| b (multiplicity two) | $[-\frac{1}{16}, \frac{1}{2}]$ |
| $\frac{1}{2}$ | $\{\frac{1}{2}\}$ |
| $-2b$ | $[-1, \frac{1}{8}]$ |
| $\frac{1 + \sqrt{1 + 16b}}{4} = \begin{cases} -\frac{\sin \theta}{2} & (-\frac{\pi}{2} \leq \theta \leq -\frac{\pi}{6}) \\ \frac{1 + \sin \theta}{2} & (-\frac{\pi}{6} \leq \theta \leq \frac{\pi}{2}) \end{cases}$ | $[\frac{1}{4}, 1]$ |
| $\frac{1 - \sqrt{1 + 16b}}{4} = \begin{cases} \frac{1 + \sin \theta}{2} & (-\frac{\pi}{2} \leq \theta \leq -\frac{\pi}{6}) \\ -\frac{\sin \theta}{2} & (-\frac{\pi}{6} \leq \theta \leq \frac{\pi}{2}) \end{cases}$ | $[-\frac{1}{2}, \frac{1}{4}]$ |

We see that unless $b = \frac{1}{2}$, i.e., $\theta = \frac{\pi}{2}$, 1 is the unique peripheral eigenvalue, and it has multiplicity 1. By Theorem 2.5, the cascade approximants then do indeed converge to the scaling function. If $\theta = \frac{\pi}{2}$, then 1 is an eigenvalue of multiplicity 2 and -1 is an eigenvalue of multiplicity 1, and by Theorem 2.5 the cascades do not converge in L^2 -norm. The slow convergence near $\theta = \frac{\pi}{2}$ can be explained by the fact that R then has eigenvalues differing from ± 1 by $O\left((\theta - \frac{\pi}{2})^2\right)$, so

the approximation after n steps is like $\left(1 - \text{const.} \left(\theta - \frac{\pi}{2}\right)^2\right)^n$. This slow rate of convergence near $\frac{\pi}{2}$ is also clear from Figure 6(a)–(1). We should emphasize that the semiregular layers displayed by these pictures dissolve more and more when doing further iterations and a plot after 1000 iterations with a resolution of 2^{-10} shows virtually no discernible small-scale regularity. This is shown in Figures 7–9, and let us explain how these were produced. First note that by (2.9), the value of $(M\psi)(x)$ at a point $x \in 2^{-N}\mathbb{Z}$ only depends on the values of $\psi(y)$ at points $y \in 2^{-(N-1)}\mathbb{Z}$, and hence the exact values $\psi^{(n)}$ for $x \in 2^{-N}\mathbb{Z}$ can be determined exactly by an iterated matrix scheme. Also the jumps of $\psi^{(n)}$ at the points in $2^{-N}\mathbb{Z}$ can be determined exactly by the following cascade scheme with fixed N (e.g., $N = 10$): We define

$$x_n = n \cdot 2^{-N}, \quad n = 0, 1, \dots, 3 \cdot 2^N,$$

and define

$$\psi_+^{(m)}(n) = \lim_{x \rightarrow n \cdot 2^{-N}+} \psi^{(m)}(x), \quad \psi_-^{(m)}(n) = \lim_{x \rightarrow n \cdot 2^{-N}-} \psi^{(m)}(x),$$

where

$$\psi^{(0)}(x) = \begin{cases} 1 & 0 \leq x < 1, \\ 0 & 1 \leq x \leq 3, \end{cases}$$

and

$$\psi^{(m)}(x) = \left(M^m \psi^{(0)}\right)(x) = M \psi^{(m-1)}(x),$$

and

$$(3.3) \quad M\psi(x) = \sqrt{2} \sum_{k=0}^3 a_k \psi(2x - k).$$

We start the recursion with

$$\psi_+^{(0)}(n) = \begin{cases} 1 & \text{for } n = 0, 1, \dots, 2^N - 1, \\ 0 & \text{for } n = 2^N, 2^N + 1, \dots, 3 \cdot 2^N, \end{cases}$$

and

$$\psi_-^{(0)}(n) = \begin{cases} 0 & \text{for } n = 0, \\ 1 & \text{for } n = 1, 2, \dots, 2^N, \\ 0 & \text{for } n = 2^N + 1, \dots, 3 \cdot 2^N, \end{cases}$$

and from (3.3),

$$(3.4) \quad \begin{cases} \psi_+^{(m)}(n) = \sqrt{2} \sum_{k=0}^3 a_k \psi_+^{(m-1)}(2n - k \cdot 2^N), \\ \psi_-^{(m)}(n) = \sqrt{2} \sum_{k=0}^3 a_k \psi_-^{(m-1)}(2n - k \cdot 2^N), \end{cases}$$

where we use the convention that

$$\psi_{\pm}^{(m-1)}(l) = 0$$

if $l < 0$ or $l > 3 \cdot 2^N$.

In Figures 7–9, we have used this algorithm to plot $\psi_+^{(1000)}(n)$, $\psi_-^{(1000)}(n)$, $(\psi_+^{(1000)} - \psi_-^{(1000)})(n)$ for $N = 10$, $\theta = 9\pi/20$. If 1000 is replaced by 100 the large-scale plots of $\psi_-^{(100)}$ and $\psi_+^{(100)}$ look similar, but $\left\| \psi_+^{(100)} - \psi_-^{(100)} \right\|_{\infty}$ is much larger

than $\left\| \psi_+^{(1000)} - \psi_+^{(1000)} \right\|_\infty < 5 \cdot 10^{-6}$. Thus the plot shows an amazing amount of local continuity even though the larger-scale behaviour is quite irregular. However, the scaling function at $\theta = 9\pi/20$ is indeed discontinuous by the discussion below. This apparent non-compatibility is explained in detail in the Appendix.

Remark 3.1. In [BEJ99], we noted that there are two neighborhoods on the circle, one near $\theta_1 = \frac{7\pi}{6}$, and a symmetric one near $\theta_2 = -\frac{\pi}{6}$, such that each of the scaling functions $\varphi^{(\theta)}$ has $x \mapsto \varphi^{(\theta)}(x)$ continuous when θ is in the union of the two neighborhoods. On the other hand if $0 < \theta < \frac{\pi}{2}$, then $\sqrt{2}a_3^{(\theta)} > 1$ so at the right-hand endpoint x of each of the dyadic partitions, we will have $(M^n \psi^{(0)})(x) \xrightarrow{n \rightarrow \infty} \infty$. It then follows from [Dau92, Proposition 6.5.2 and footnote 9] or [DaLa92] that $x \mapsto \varphi^{(\theta)}(x)$ cannot be continuous for θ in the first quarter circle. By the symmetry $\varphi^{(\pi-\theta)}(x) = \varphi^{(\theta)}(3-x)$, it is then also not continuous in the second quarter of the θ -circle.

The question of continuity of $\varphi^{(\theta)}$ has been considered in even more detail for our examples in the papers [CoHe92], [CoHe94, Section 4.8.1], [Wan95], [Wan96]. They use the real coefficients

$$(3.5) \quad c_n = \sqrt{2}a_n,$$

which then satisfy

$$(3.6) \quad \begin{aligned} \sum_k c_k c_{k+2l} &= 2\delta_l, \\ \sum_k c_k &= 2, \end{aligned}$$

or, equivalently,

$$(3.7) \quad c_0 + c_2 = 1 = c_1 + c_3,$$

so the scaling operator M in (2.9) becomes

$$(3.8) \quad (M\psi)(x) = \sum_{k=0}^3 c_k \psi(2x - k),$$

where

$$(3.9) \quad \begin{aligned} c_0 &= \frac{1}{2}(1 - \cos \theta + \sin \theta), \\ c_1 &= \frac{1}{2}(1 - \cos \theta - \sin \theta), \\ c_2 &= \frac{1}{2}(1 + \cos \theta - \sin \theta), \\ c_3 &= \frac{1}{2}(1 + \cos \theta + \sin \theta); \end{aligned}$$

see an illustration in Figure 2. In [Wan96, Proposition 2.1], it is stated that if there is a continuous scaling function, then $|c_0| < 1$ and $|c_3| < 1$, that is, $\pi < \theta < 2\pi$. Since our movie reel shows that the scaling function clearly is discontinuous for $0 \leq \theta \leq \frac{\pi}{2}$ and then by symmetry for $0 \leq \theta \leq \pi$, this is consistent with the movie reel. (It is still an open question, for $0 < \theta < \pi$, as to “how discontinuous” $x \mapsto \varphi^{(\theta)}(x)$ then is. Based on graphics, and analogies (see Section 4) to iterated function systems, it is likely that the discontinuous cases have interesting fractal

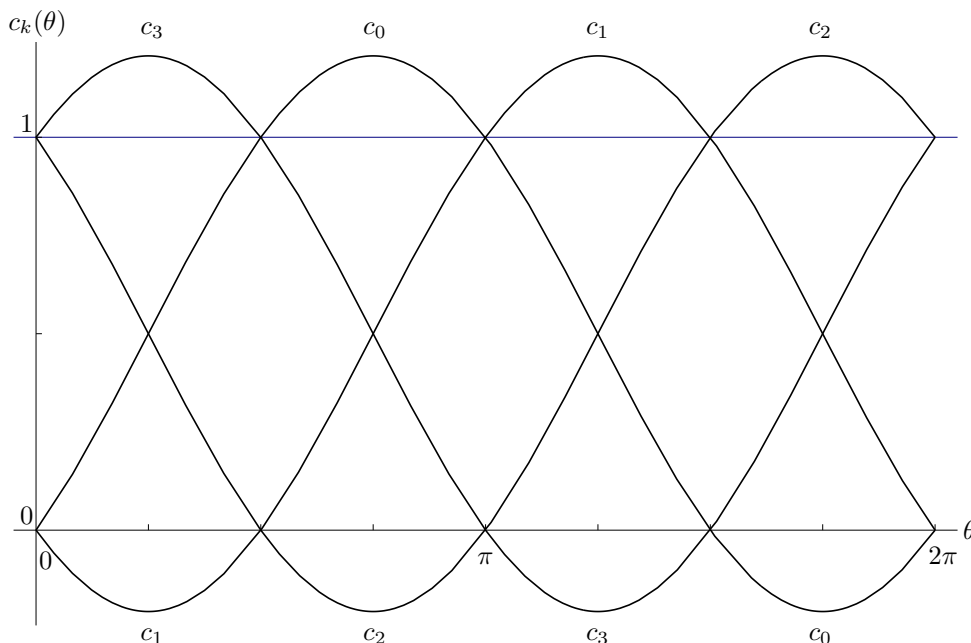


FIGURE 2. $c_k(\theta) = \sqrt{2}a_k^{(\theta)}$, $k = 0, 1, 2, 3$. See (3.9).

structure, but that will be postponed to a later paper.) On the other hand, the comment to Figure 4.3 in [CoHe94, page 193] indicates that the condition $\theta \in \langle \pi, 3\pi/2 \rangle \cup \langle 3\pi/2, 2\pi \rangle$ is necessary and sufficient for a continuous scaling function. This is indeed consistent with our movie reel, but the reel shows extremely singular behaviour of the scaling function at some dyadic rationals even in the domain of continuity. The effect is most pronounced for Figures 3(b) and 3(j). We refer to the papers above and [DaLa92] for the actual methods used to establish continuity.

The flip symmetry of the example above holds more generally whenever the Ruelle matrix is defined from a low-pass filter function $m_0(z) = a_0 + a_1z + \cdots + a_Nz^N$ with real coefficients a_i . Specifically let R denote the associated Ruelle operator. We then have

Lemma 3.2.

$$(R\xi)(z^{-1}) = (R\xi^\vee)(z),$$

where $\xi^\vee(z) := \xi(z^{-1})$, $\xi \in \mathbb{C}[z, z^{-1}]$, $z \in \mathbb{T}$.

Proof. From the definition of R we have

$$\begin{aligned}
 (R\xi)(z^{-1}) &= \frac{1}{2} \sum_{w^2=z^{-1}} |m_0(w)|^2 \xi(w) \\
 &= \frac{1}{2} \sum_{w^2=z} |m_0(w^{-1})|^2 \xi(w^{-1}) \\
 &= \frac{1}{2} \sum_{w^2=z} |\overline{m_0(w)}|^2 \xi^\vee(w) \\
 &= R(\xi^\vee)(z),
 \end{aligned}$$

where we used the reality assumption in the form $m_0(w^{-1}) = \overline{m_0(w)}$. \square

We now turn to the graphics which illustrate the cascade approximation (2.9)–(2.10). It follows from (2.1) and [Dau92, pp. 205–6] that this can be based on the

matrix $\begin{pmatrix} a_0 & 0 & 0 & 0 \\ a_2 & \boxed{a_1 \ a_0} & 0 \\ 0 & \boxed{a_3 \ a_2} & a_1 \\ 0 & 0 & 0 & a_3 \end{pmatrix}$, or alternatively on just the framed 2×2 matrix $\begin{pmatrix} a_1 & a_0 \\ a_3 & a_2 \end{pmatrix}$.

The so-called cascade algorithm [Dau92, p. 205] starts with initial points on an integral grid, and the n 'th step fills in points on $2^{-n}\mathbb{Z}$ places according to matrix multiplication and use of the weights from the respective matrix entries. We include more technical points in the captions of the pictures, and in the Appendix.

Remark 3.3. While the results above concern primarily orthogonality and L^2 -cascade approximation, there is a direct connection between the L^2 -theory and pointwise features of the approximation, as is pointed out in [Dau92, p. 204]. Proposition 6.5.2 in [Dau92] makes that explicit when the *a priori* assumption is made that $\varphi^{(\theta)}(x)$ is continuous in x . As noted in Remark 3.1 and Section 4, and in [BEJ99], such continuity is only known when θ is restricted to certain subintervals of $(-\pi, \pi]$, and the pictures serve to illustrate the features when θ is in the complement of the “good” regions.

We stress that the fast algorithm used for some of the graphics (see the Appendix) does in fact depend on the orthogonality of the family $\{\varphi^{(\theta)}(x - n) \mid n \in \mathbb{Z}\}$ in $L^2(\mathbb{R})$, and this orthogonality we verified in [BEJ99] to be satisfied for all values of θ except for $\theta = \frac{\pi}{2}$. The significance of this orthogonality is also directly related to the assumption made in Theorem 2.5(ii), i.e., (2.24), on the starting function $\psi^{(0)}$ for the cascade approximation (2.12), but the fast algorithm of the Appendix is different from (2.9)–(2.12). Details on the comparison of the two are to be found in the Appendix.

LIST OF FIGURES

| | | |
|---|--|-------|
| 1 | Eigenvalues $\lambda(\theta)$ of R | 16 |
| 2 | $c_k(\theta) = \sqrt{2}a_k^{(\theta)}$, $k = 0, 1, 2, 3$ (see comments below) | 19 |
| 3 | Scaling function “movie reel”, θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in twenty-one frames (see comments below) | 22–25 |
| 4 | The $m \rightarrow \infty$ limit of $\psi^{(m)}(x)$ for $\theta = -\frac{9\pi}{20}$ | 26 |
| 5 | Initial function $\psi^{(0)}(x) = \chi_{[0,1]}(x)$ for the cascade series (Haar scaling function) | 27 |

| | | |
|----|---|-------|
| 6 | Cascade stages of scaling function, $\theta = \frac{9\pi}{20}$ (see comments below) | 28–30 |
| 7 | $\psi_+^{(1000)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ | 31 |
| 8 | $\psi_-^{(1000)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ | 31 |
| 9 | $\left(\psi_+^{(1000)}(n) - \psi_-^{(1000)}(n)\right)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ | 32 |
| 10 | $\psi_{\pm}^{(\infty)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ | 32 |

In Figure 6(a)–(l), we study the cascade approximation at $\theta = \frac{9\pi}{20}$ with the Haar function as starting point. Special attention will be given to the asymptotic properties of the jumps at certain dyadic rationals. The rightmost term in the sum expansion for $M^n \psi^{(0)}(x) = \sum_i h_i(n) \psi^{(0)}(2^n x - i)$ is

$$h_{N_n}(n) \psi^{(0)}(2^n x - N_n),$$

where $N_n = 3(2Hn - 1 + \dots + 2 + 1) = 3 \cdot (2Hn - 1)$, and where $\sqrt{2}a_3 > 1$, $h_{N_n}(n) = 2^{\frac{n}{2}} a_3^n \xrightarrow{n \rightarrow \infty} \infty$, and each of the previous subpartition highpoints $h_{N_n - k3}(n)$ contains a term with a factor $2^{\frac{n}{2}} a_3^{n-k} \xrightarrow{n \rightarrow \infty} \infty$, $k = 0, 1, 2, \dots$. For example, note that $\psi^{(n)} = M^n \psi^{(0)}$ is supported in $[0, x_n)$, where $x_n = 3 - 2^{1-n}$, and

$$\lim_{x \rightarrow x_n^-} \psi^{(n)}(x) = \left(\sqrt{2}a_3\right)^n \xrightarrow{n \rightarrow \infty} \infty.$$

Figure 6(a): $\psi^{(1)}(x) = M\psi^{(0)}(x) = \sqrt{2} \sum_i a_i \psi^{(0)}(2x - i)$; cf. (2.9). The a_i 's are given by (3.2).

Figure 6(b): $\psi^{(2)}(x) = M^2 \psi^{(0)}(x) = 2 \sum_k \sum_i a_i a_{k-2i} \psi^{(0)}(4x - k) = \sum_k h_k(2) \psi^{(0)}(4x - k)$.

Figure 6(c): $\psi^{(3)}(x) = M^3 \psi^{(0)}(x) = 2^{\frac{3}{2}} \sum_k \sum_{ij} a_i a_j a_{k-4i-2j} \psi^{(0)}(8x - k) = \sum_k h_k(3) \psi^{(0)}(8x - k)$.

Figure 6(d): $\psi^{(4)}(x) = M^4 \psi^{(0)}(x) = \sum_k h_k(4) \psi^{(0)}(2^4 x - k)$, where the height of the rightmost column is $h_{45}(4) = (\sqrt{2}a_3)^4$, and further high points at $h_{42}(4)$, $h_{39}(4)$, $h_{36}(4)$, \dots , corresponding to rightmost “bumps” in the subpartition intervals.

\vdots

Figure 6(l): The $n = 12$ case of $\psi^{(n)}(x) = M^n \psi^{(0)}(x) = \sum_k h_k(n) \psi^{(0)}(2^n x - k)$, where $2^{-n} \sum_k h_k(n)^2 = \|M^n \psi^{(0)}\|_2^2 \|\psi^{(0)}\|_2^2 = 1$ by (2.81), despite the divergence at the high rightmost “bumps”, $n \rightarrow \infty$.

Additional comments on the Figures: The numerical sizes of the y -coordinates of the cusp points in Figure 3 are computed and displayed in the Appendix. Note that figures 3(b)–3(j) represent continuous functions, while all other figures represent discontinuous functions [CoHe94, p. 193]. See more details of Figure 3(b) in Figure 4.

See <http://cm.bell-labs.com/who/wim/cascade/> for more pictures, but without emphasis of the singularities at dyadic points.

In Figure 2, note that everywhere except at the four points $\theta = k\frac{\pi}{2}$, $k = 0, 1, 2, 3$, we have precisely three positive a_i 's and one negative one. The significance of that is discussed in Section 4 below.

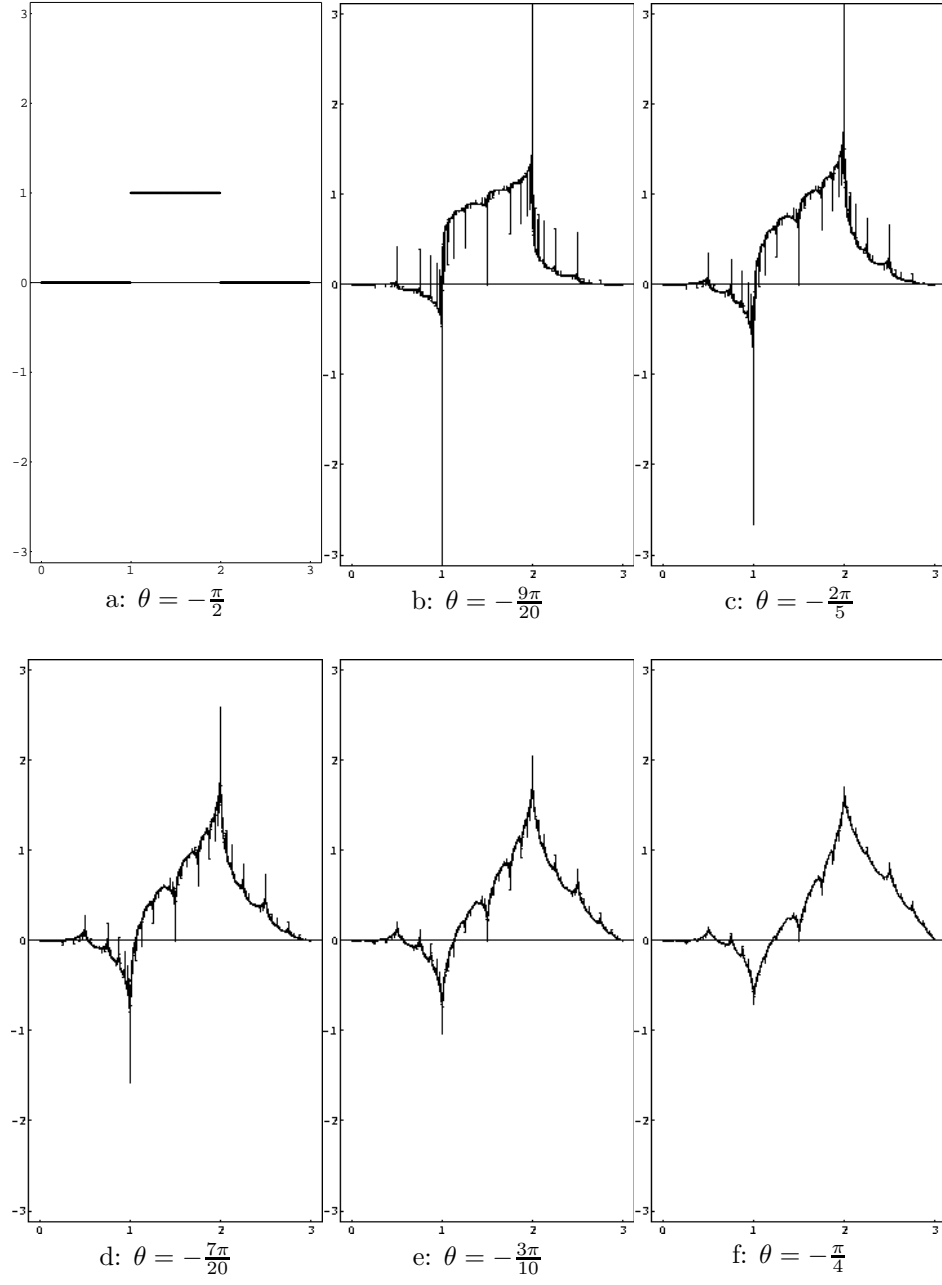


FIGURE 3. Scaling function “movie reel”, θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in twenty-one frames: Frames a–f

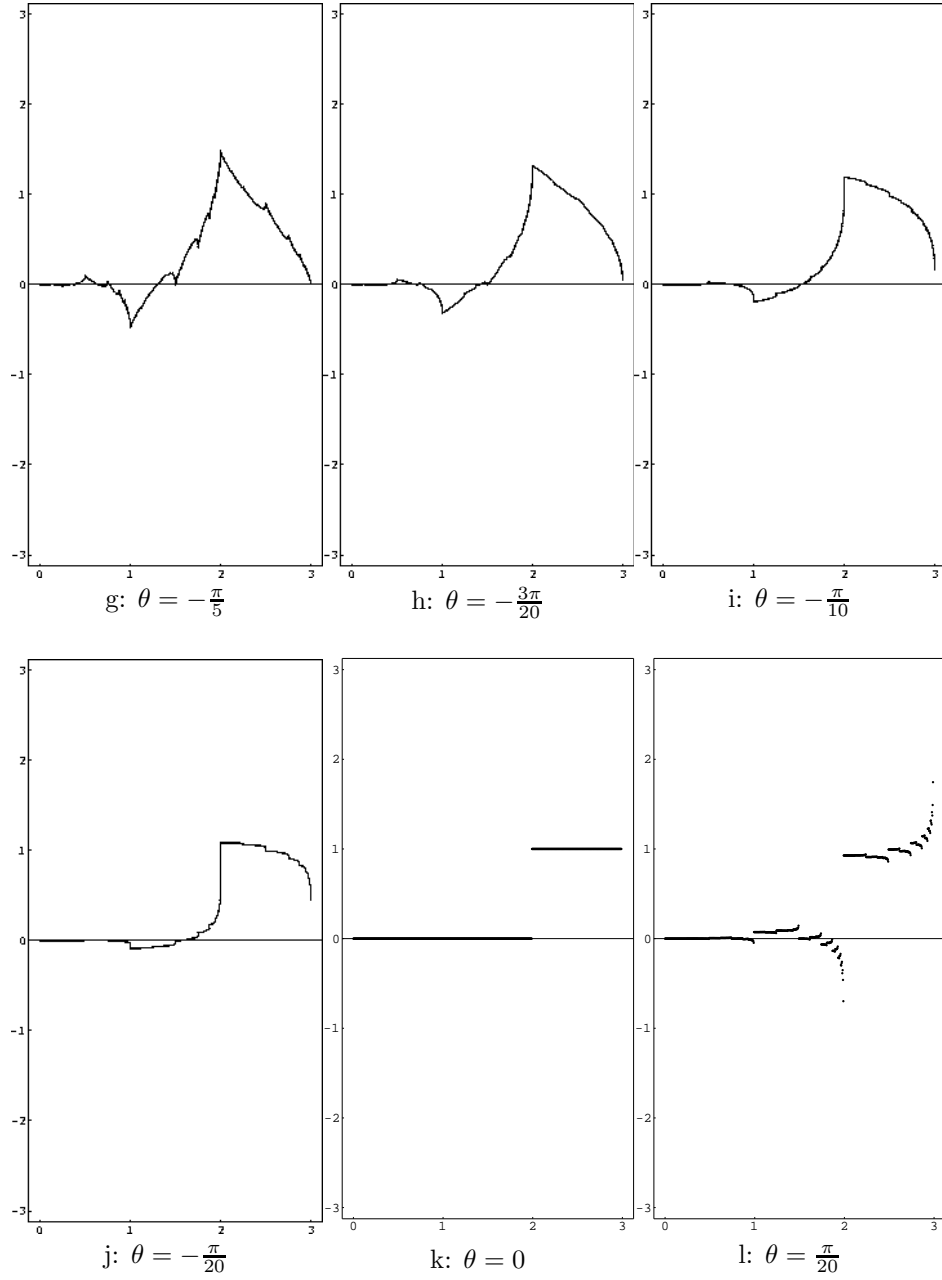


FIGURE 3. Scaling function "movie reel", θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in twenty-one frames: Frames g-l

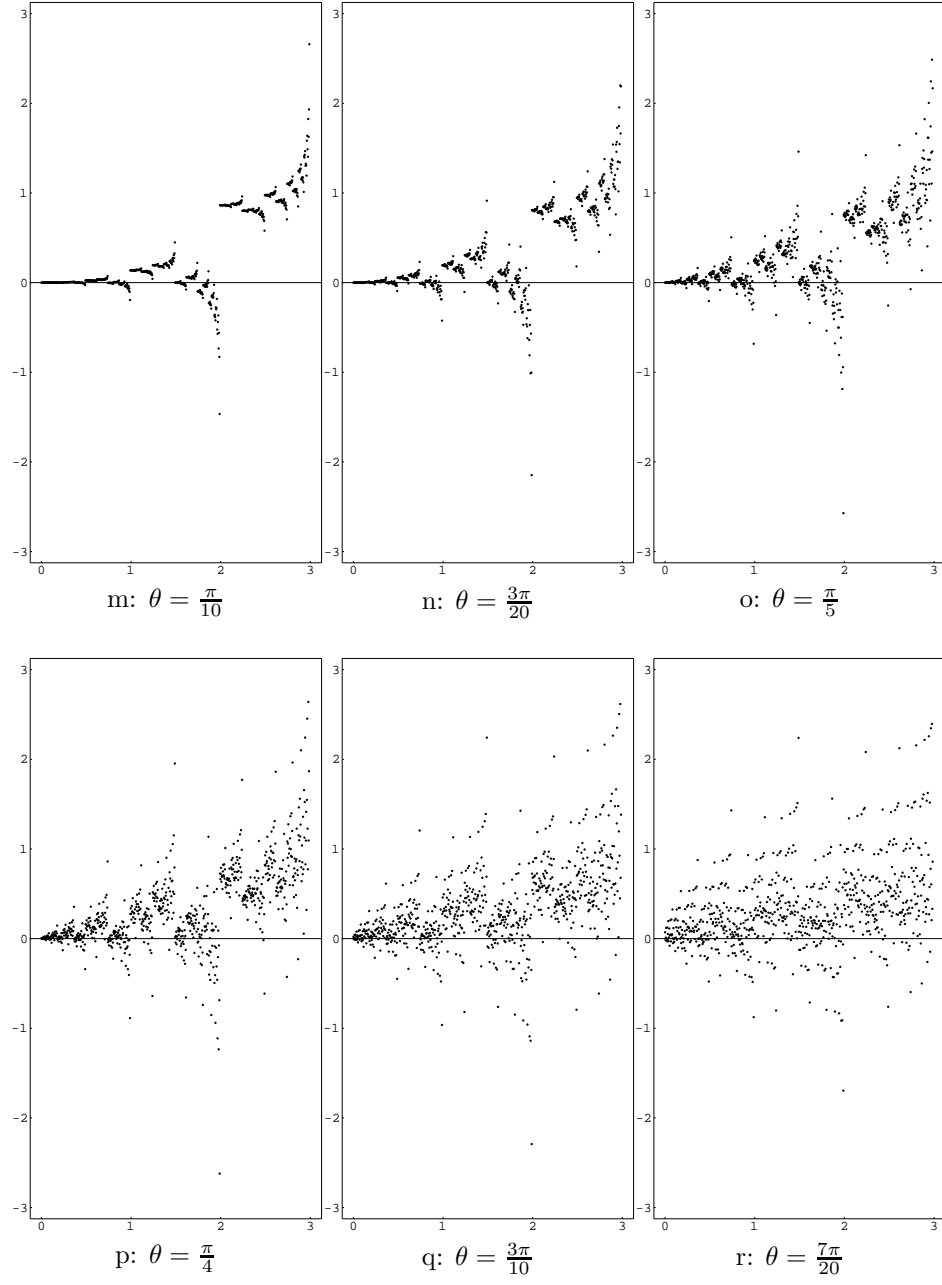


FIGURE 3. Scaling function "movie reel", θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in twenty-one frames: Frames m-r

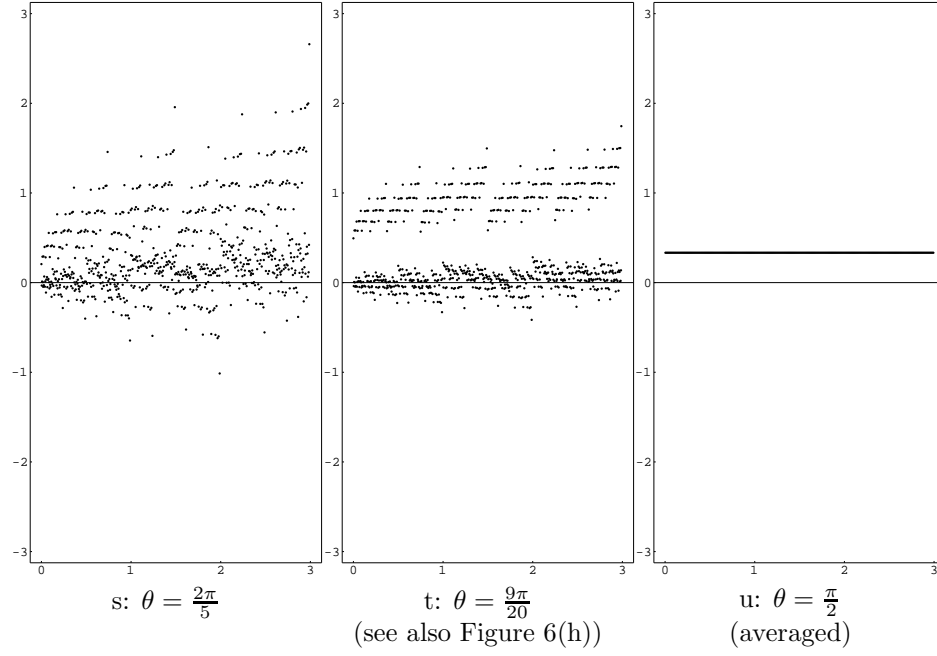


FIGURE 3. Scaling function “movie reel”, θ from $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ in twenty-one frames: Frames s–u

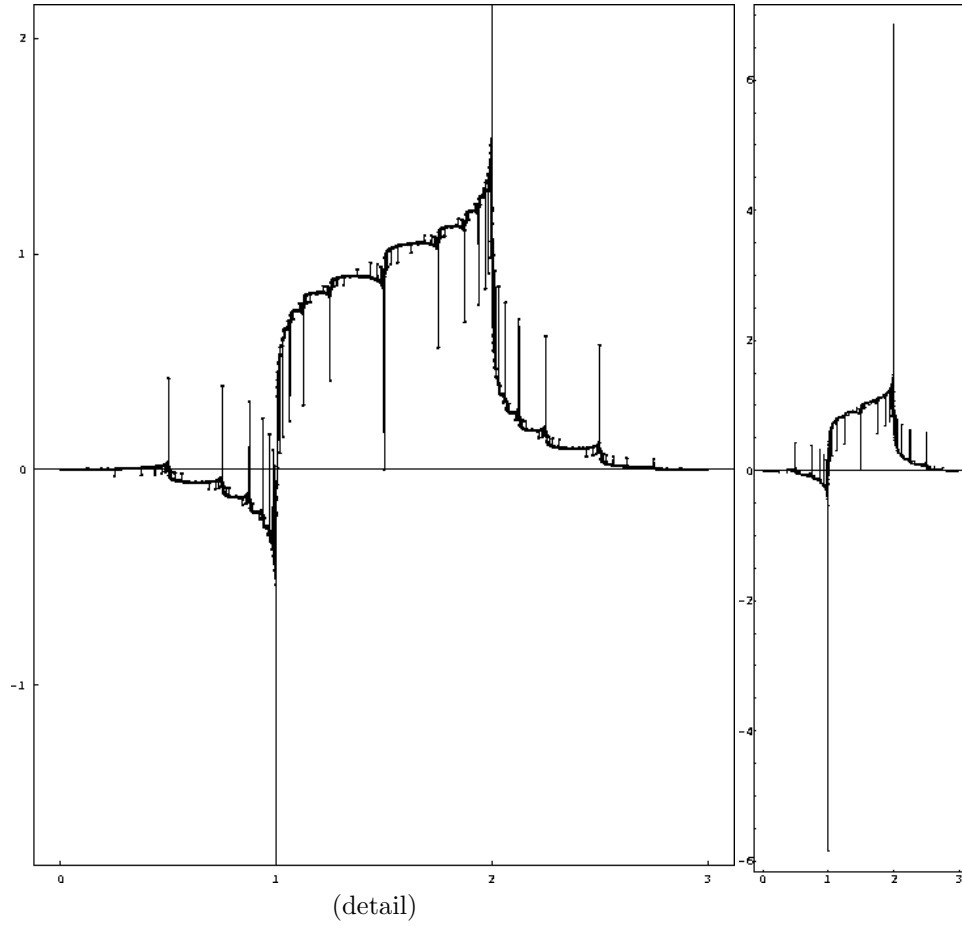


FIGURE 4. The $m \rightarrow \infty$ limit of $\psi^{(m)}(x)$ for $\theta = -\frac{9\pi}{20}$

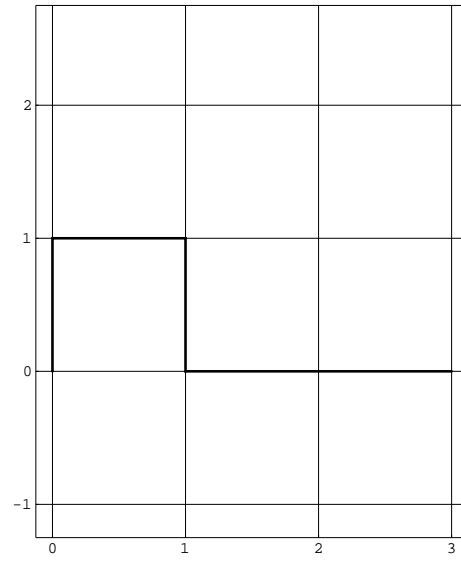


FIGURE 5. Initial function $\psi^{(0)}(x) = \chi_{[0,1]}(x)$ for the cascade series (Haar scaling function)

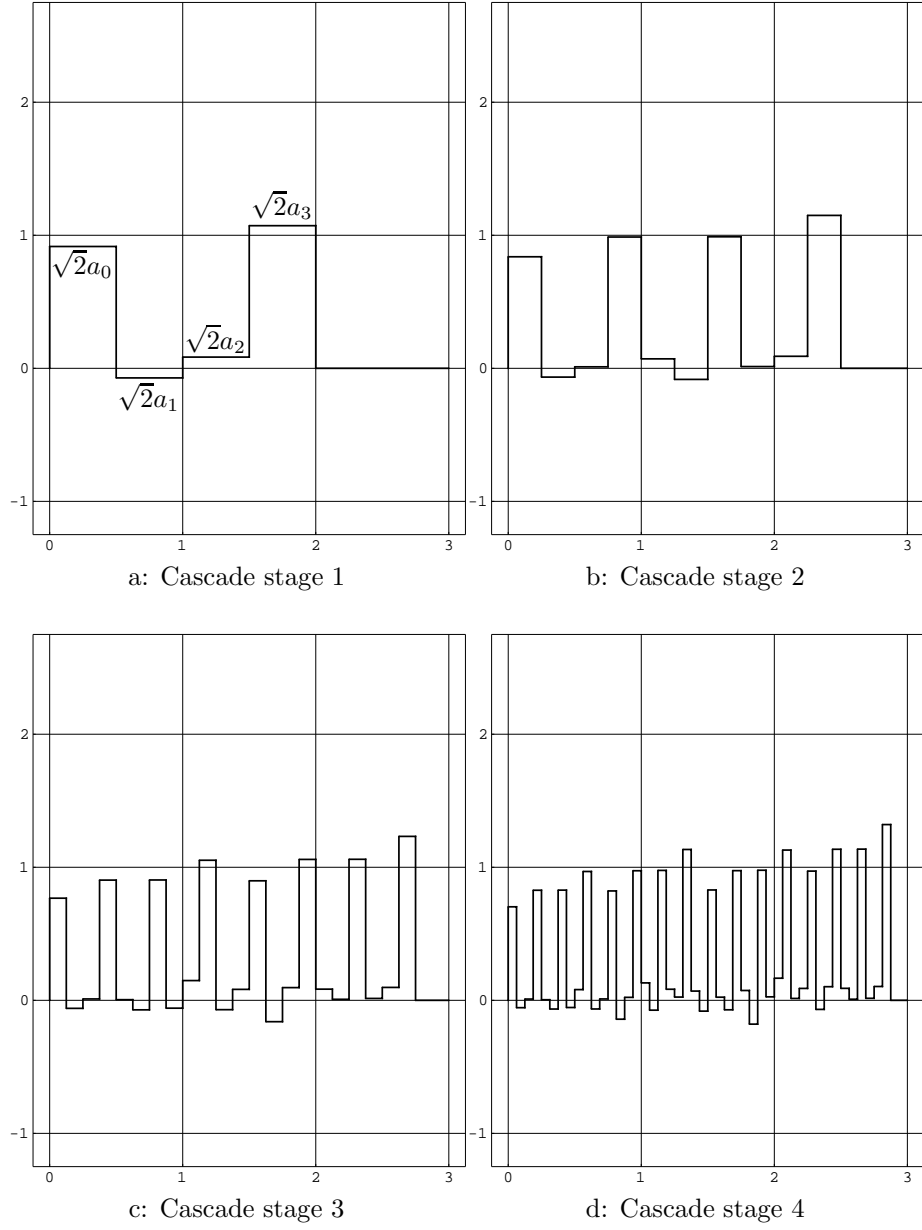
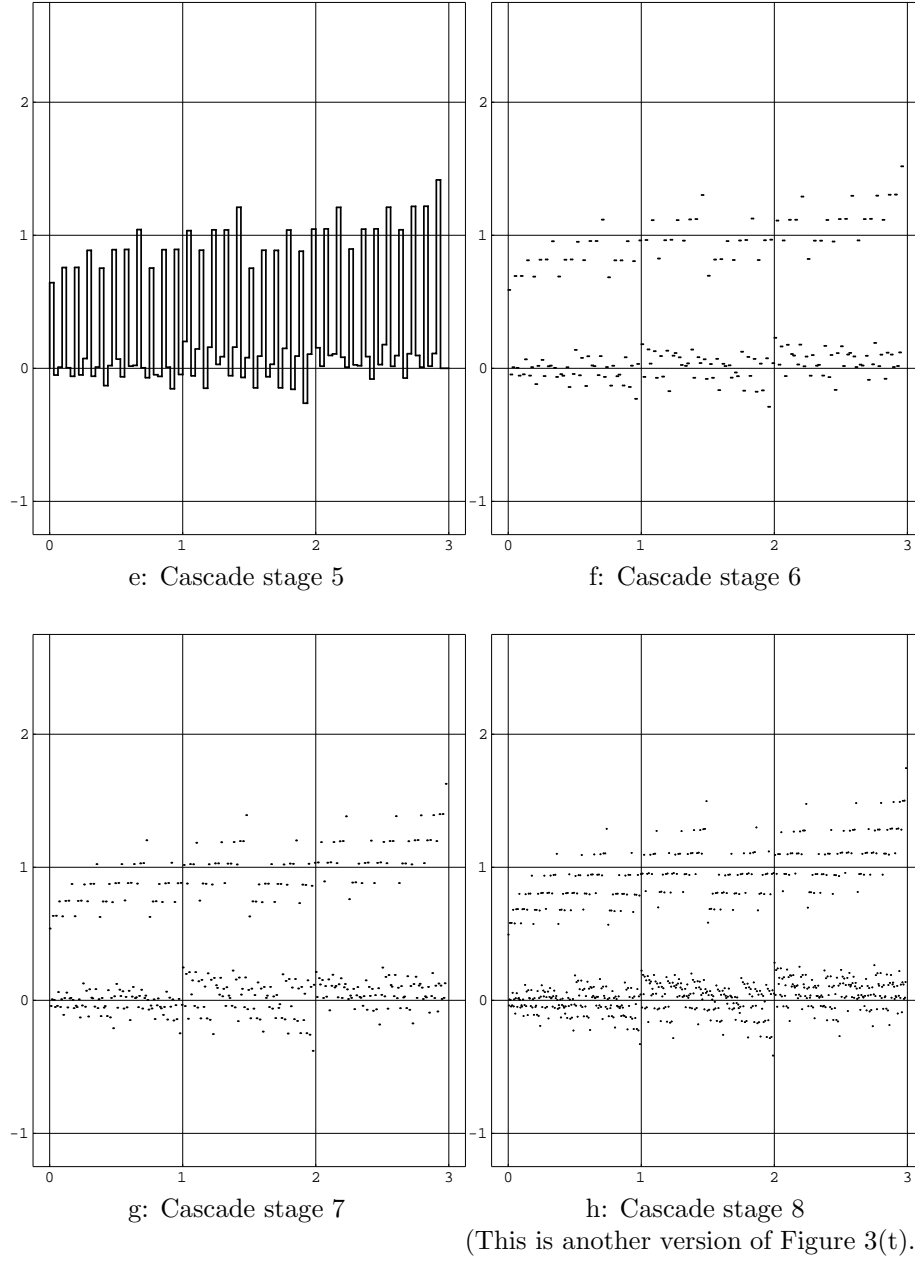


FIGURE 6. Cascade stages of scaling function, $\theta = \frac{9\pi}{20}$: Stages 1–4


 FIGURE 6. Cascade stages of scaling function, $\theta = \frac{9\pi}{20}$: Stages 5–8

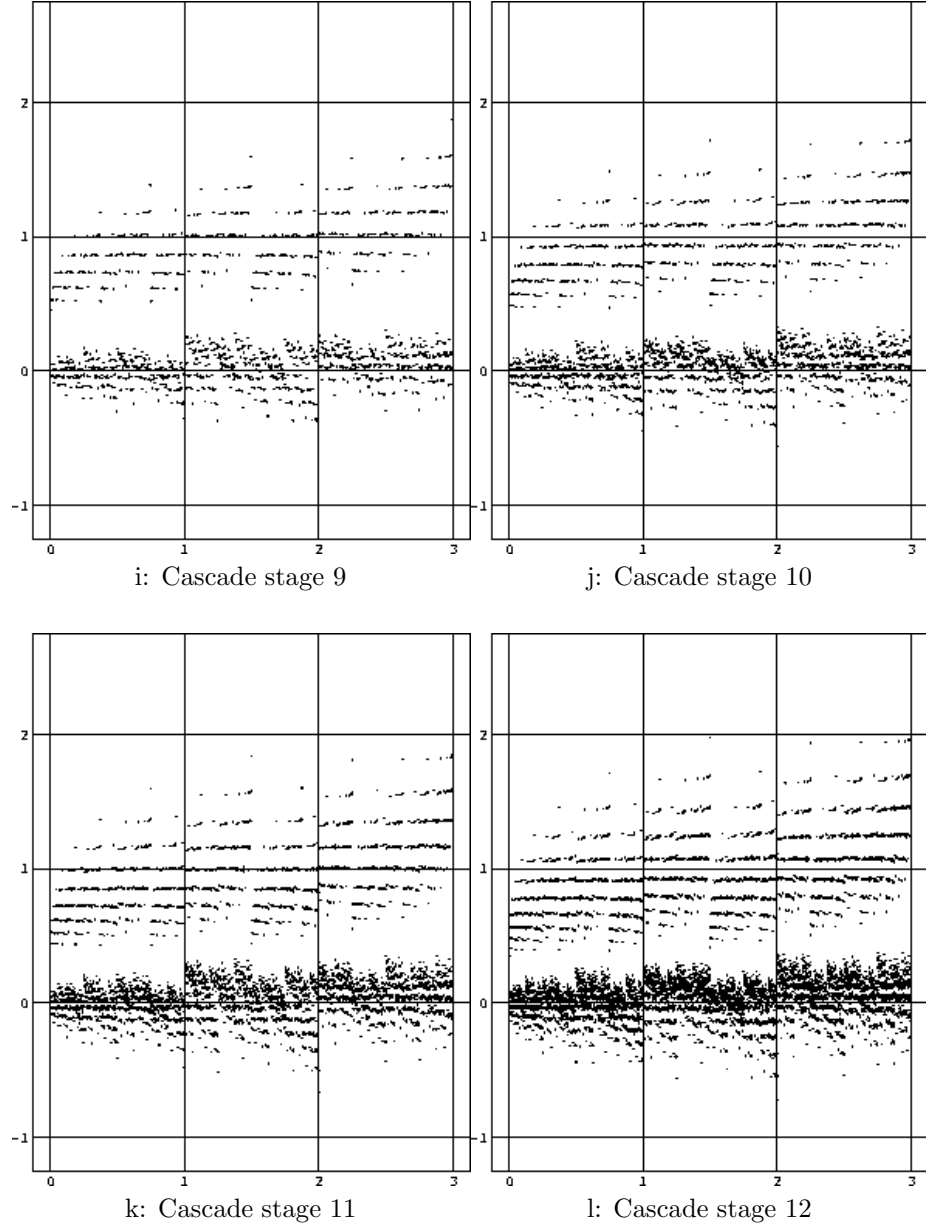


FIGURE 6. Cascade stages of scaling function, $\theta = \frac{9\pi}{20}$: Stages 9–12

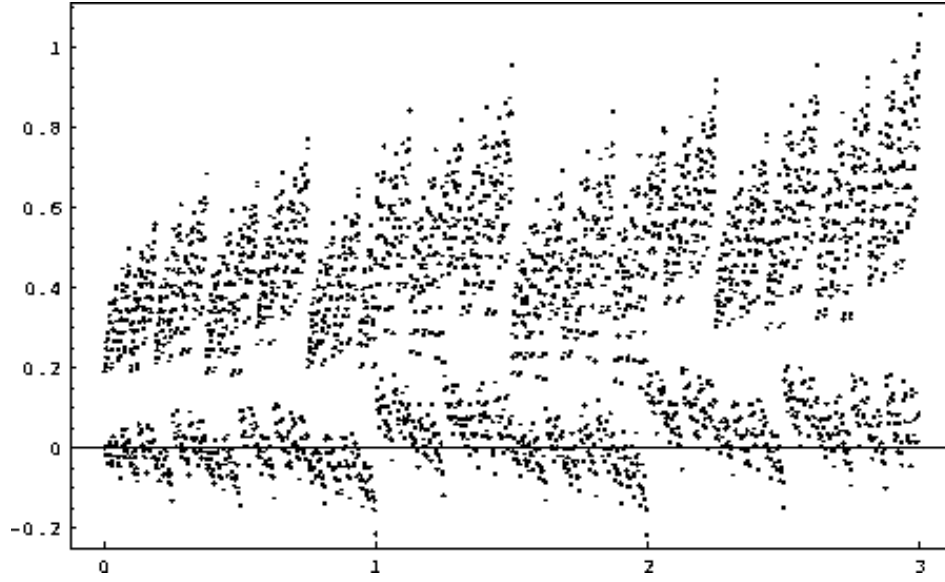


FIGURE 7. The figure above shows $\psi_+^{(1000)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ ($\theta = \frac{9\pi}{20}$).

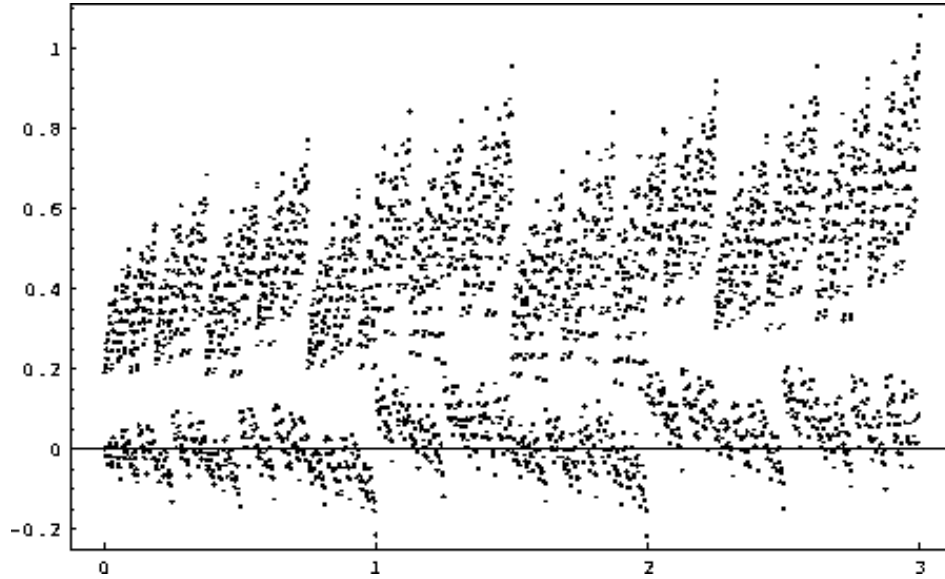


FIGURE 8. The figure above shows $\psi_-^{(1000)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ ($\theta = \frac{9\pi}{20}$).

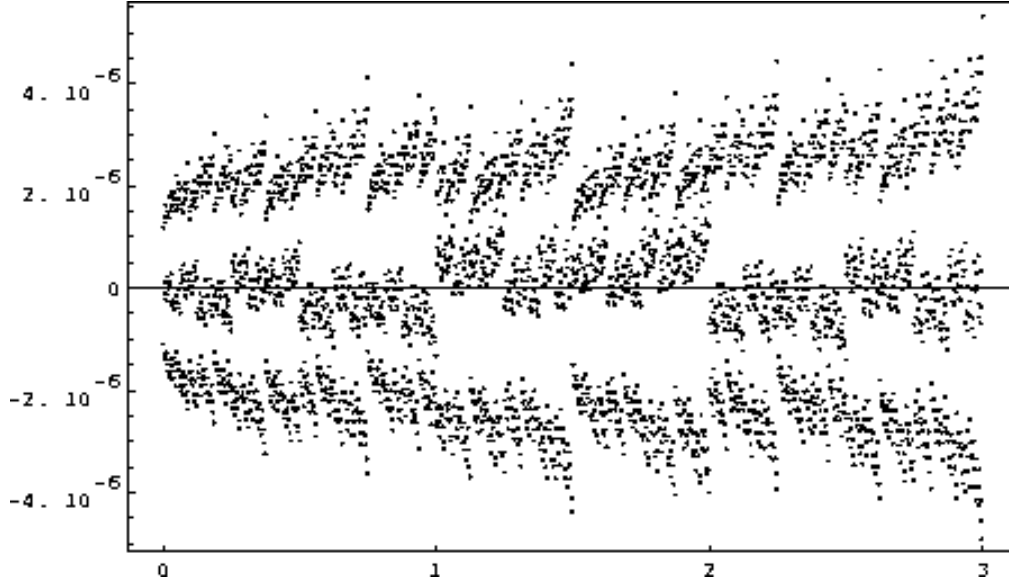


FIGURE 9. The figure above shows $\left(\psi_+^{(1000)}(n) - \psi_-^{(1000)}(n)\right)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$. Note the small scale at the y -axis, showing $\left|\psi_+^{(1000)}(n) - \psi_-^{(1000)}(n)\right| < \frac{5}{1000000}$. See the Appendix for a discussion of this convergence of the differences to zero.

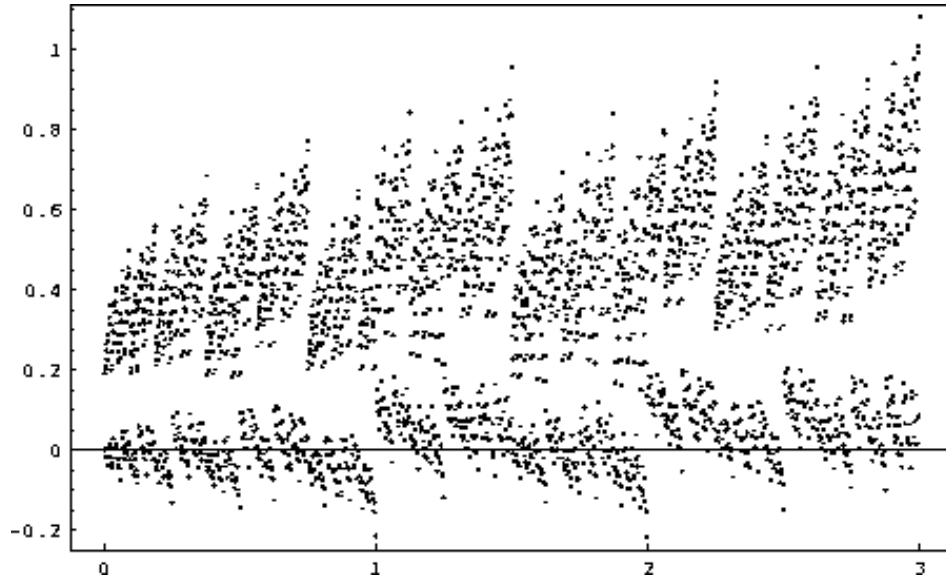


FIGURE 10. The figure above shows $\psi_{\pm}^{(\infty)}(n)$ at $n = 0, 1, \dots, 3 \cdot 2^{10}$, $x_n = n \cdot 2^{-10}$ ($\theta = \frac{9\pi}{20}$).

4. CONCLUSIONS

While it may be difficult to discern an overall pattern in the computer-generated output from the cascades, the expectation is that there are domains of starting points (functions or coefficients) which lead to “nice” limit functions, $L^2(\mathbb{R})$ or continuous, while the other extreme ones lead to fractal-like pictures. This is based on two analogies, in addition to the existing results in [DaLa92], [CoHe92], [CoHe94], [Wan95], [Wan96]:

(i) First, we think of (2.1) as a version of an iterated function system in the sense of Hutchinson [Hut81]. Hutchinson considers Borel mappings f_1, \dots, f_n on a complete metric space X , probabilities p_1, \dots, p_n , $p_i > 0$, $\sum_i p_i = 1$. Such a system defines a dynamical system $x_0 \rightarrow x_1 \rightarrow \dots$ in X where

$$x_{i+1} = f_{\alpha_i}(x_i)$$

and where the indices α_i are chosen randomly for each i , with probability p_{α_i} . For Borel probability measures μ on X , set $M_i(\mu) = \mu \circ f_i^{-1}$, i.e., $M_i(\mu)(E) = \mu(f_i^{-1}(E))$, where $E \subset X$ is a Borel set, and $f_i^{-1}(E) = \{x \in X \mid f_i(x) \in E\}$. A fixed point for the system is a measure μ such that

$$(4.1) \quad \mu = \sum_i p_i M_i(\mu),$$

and we think of (2.1) as a version of this, but of course, in (2.1) we do not necessarily have the coefficients a_i positive, and we do not impose the same normalization. Nonetheless, we may take $X = \mathbb{R}$, $f_i(x) := \frac{1}{2}(x + i)$, and $p_i := \frac{1}{\sqrt{2}}a_i$, $i = 0, 1, \dots, N$. (Figure 2 shows examples with four coefficients, where generically three are positive and one is negative.) Hutchinson’s theorem for the general version of (4.1) yields existence and uniqueness of μ provided the probabilistic assumptions hold, and the mappings f_i have contractive Lipschitz constants. Even in the wavelet setting, one may ask for a signed measure μ solving (4.1) and make the distinction between solutions μ which are absolutely continuous with respect to Lebesgue measure, vs. the measures μ which are singular. In the first case, there is a Radon-Nikodym derivative φ which may be viewed as a solution to the original problem (2.1).

(ii) The second analogy is to a problem studied by P. Erdős [Erd40] for an iterated function system on

$$(4.2) \quad \begin{aligned} X &= [0, 1], & 0 < \lambda < 1, \\ f_0(x) &= \lambda x, \\ f_1(x) &= \lambda x + 1 - \lambda, & \text{and} \\ p_0 &= p_1 = \frac{1}{2}. \end{aligned}$$

The corresponding problem (4.1) and (4.2) leads to a probability measure μ_λ , and Erdős showed that μ_λ is either absolutely continuous, or else totally singular. Total singularity means that there is a subset $E \subset X (= [0, 1])$ such that $\mu_\lambda(E) = 1$ and E is of zero Lebesgue measure. This shows up in computer output as fractal-like appearance. It is known that if $\lambda < \frac{1}{2}$, then the support of μ_λ is a Cantor set of Hausdorff dimension $-\frac{\ln 2}{\ln \lambda}$, while the case $\lambda > \frac{1}{2}$ is not fully understood, the expectation being that μ_λ will be more regular (or “less fractal”) if λ is closer to 1.

The case $\lambda = \frac{1}{2}$ in the Erdős construction is clear, of course, and yields $d\mu_{\frac{1}{2}}(x) = \varphi(x) dx$ where $\varphi = \psi^{(0)}$ is precisely the Haar function from (2.13) above, and dx is Lebesgue measure, i.e., φ satisfies

$$\varphi\left(\frac{x}{2}\right) = \varphi(x) + \varphi(x-1),$$

which is also (2.1) for the Haar scaling function.

It should also be mentioned that Solomyak [Sol95] proved that for almost all λ in $(\frac{1}{2}, 1)$, the measure μ_λ does have an $L^2(\mathbb{R})$ density. An example of $\lambda > \frac{1}{2}$ when μ_λ is known to be singular is $\lambda^{-1} = (1 + \sqrt{5})/2$, the golden ratio.

APPENDIX

by Brian F. Treadway

By the method of Daubechies [Dau92, Section 6.5, pp. 204–206], the iteration (3.4) can be replaced by a “local” iteration in which the two constant values of $\psi^{(m)}(x)$ on adjacent intervals at the left of the point $x = k \cdot 2^{-(m-1)}$,

$$\begin{aligned} \psi_{2k-2}^{(m)} &:= \psi^{(m)}(x) & \text{for } x \in [(2k-2) \cdot 2^{-m}, (2k-1) \cdot 2^{-m}), \\ \psi_{2k-1}^{(m)} &:= \psi^{(m)}(x) & \text{for } x \in [(2k-1) \cdot 2^{-m}, 2k \cdot 2^{-m}), \end{aligned}$$

are determined as a linear combination of two constant values of $\psi^{(m-1)}(x)$ on adjacent intervals at the left of the same x from the previous iteration, viz.

$$\begin{aligned} \psi_{k-2}^{(m-1)} &:= \psi^{(m)}(x) & \text{for } x \in [(k-2) \cdot 2^{-(m-1)}, (k-1) \cdot 2^{-(m-1)}), \\ \psi_{k-1}^{(m-1)} &:= \psi^{(m)}(x) & \text{for } x \in [(k-1) \cdot 2^{-(m-1)}, k \cdot 2^{-(m-1)}). \end{aligned}$$

The relation is

$$(A.1) \quad \begin{pmatrix} \psi_{2k-2}^{(m)} \\ \psi_{2k-1}^{(m)} \end{pmatrix} = A \begin{pmatrix} \psi_{k-2}^{(m-1)} \\ \psi_{k-1}^{(m-1)} \end{pmatrix}$$

with

$$A = \begin{pmatrix} \sqrt{2}a_2 & \sqrt{2}a_0 \\ \sqrt{2}a_3 & \sqrt{2}a_1 \end{pmatrix},$$

where the a_i are as in (3.2). With the definition

$$(A.2) \quad \xi_{m-1+i} := \begin{pmatrix} \psi_{(k \cdot 2^i)-2}^{(m-1+i)} \\ \psi_{(k \cdot 2^i)-1}^{(m-1+i)} \end{pmatrix},$$

the relation (A.1) becomes

$$\xi_m = A\xi_{m-1}.$$

The matrix relation (A.1) is in fact a transcription of the equations (6.5.9) and (6.5.10) in [Dau92], and the reader should be warned that the relations are not merely an algebraic consequence of the cascade iteration (2.9), but depend heavily on the orthogonality of the translates of the successive iterates $\psi^{(m)}$. See Remark 3.3 for comments on this dependence on orthogonality. The relation between Daubechies’s iterates and ours is that

$$\psi_+^{(m)}(n) = \left(M^m \chi_{[0,1)}\right)(n \cdot 2^{-N}) = \psi_{n \cdot 2^{-N+m}}^{(m)},$$

at least when $m \geq M$. Thus, when Daubechies's scheme is iterated more than N times, the two methods give the same values on $\mathbb{Z}2^{-N}$. In order to get the values of φ at dyadic rationals it is necessary to switch from Daubechies's algorithm, which doubles the number of points at each iteration, to the algorithm based on (2.9), where the number of points is constant. The extreme cusps in the movie reel Figure 3(b)–(j) are therefore computed by using (2.9) with large n 's, or by computing the limit $n \rightarrow \infty$ as in (A.4) below. The connection between the two versions of the cascade algorithm is spelled out in [Dau92, p. 204], where equation (6.5.3) shows that the value of φ at the dyadic rational $n \cdot 2^{-N}$ may be approximated by $\psi_{2^{j-N}}^{(j)}$ as $j \rightarrow \infty$ where $\psi_m^{(j)} = 2^{j/2} \langle \varphi | \varphi_{-j,m} \rangle$ and $\varphi_{j,m}(x) = 2^{-j/2} \varphi(2^{-j}x - m)$. When the approximations at a given stage are written out in terms of products of the coefficients a_i in the order they arise in the iteration, it is found that the two algorithms supply these factors in precisely the opposite order.

If $\sin \theta \neq -1$, the eigenvectors of $A = \begin{pmatrix} \sqrt{2}a_2 & \sqrt{2}a_0 \\ \sqrt{2}a_3 & \sqrt{2}a_1 \end{pmatrix}$ are: for eigenvalue $\lambda_1 = 1$, $e_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$; for eigenvalue $\lambda_2 = -\sin \theta$, $e_2 = \begin{pmatrix} \sqrt{2}a_1 \\ \sqrt{2}a_2 \end{pmatrix}$. A starting vector $\xi_{m-1} = \begin{pmatrix} \psi_{k-2}^{(m-1)} \\ \psi_{k-1}^{(m-1)} \end{pmatrix}$ composed of ordinates in the $(m-1)$ 'st cascade approximant gives, by n applications of A ,

$$A^n \xi_{m-1} = \xi_{m-1+n},$$

where ξ_{m-1+n} , given by (A.2) above with $i = n$, is composed of ordinates in the $(m-1+n)$ 'th cascade approximant on two intervals at the left of the same point $x = k \cdot 2^{-(m-1)}$. Expanding the starting vector ξ_{m-1} in the two eigenvectors, we get

$$\begin{aligned} \xi_{m-1} &= \alpha_1 e_1 + \alpha_2 e_2 \\ &= \frac{\sqrt{2}a_2 \psi_{k-2}^{(m-1)} - \sqrt{2}a_1 \psi_{k-1}^{(m-1)}}{\sqrt{2}(a_2 - a_1)} e_1 \\ &\quad + \frac{-\psi_{k-2}^{(m-1)} + \psi_{k-1}^{(m-1)}}{\sqrt{2}(a_2 - a_1)} e_2. \end{aligned}$$

Thus

$$\begin{aligned} (A.3) \quad \xi_{m-1+n} &= A^n \xi_{m-1} \\ &= \lambda_1^n \alpha_1 e_1 + \lambda_2^n \alpha_2 e_2 \\ &= \frac{\sqrt{2}a_2 \psi_{k-2}^{(m-1)} - \sqrt{2}a_1 \psi_{k-1}^{(m-1)}}{\sqrt{2}(a_2 - a_1)} e_1 \\ &\quad + (-\sin \theta)^n \frac{-\psi_{k-2}^{(m-1)} + \psi_{k-1}^{(m-1)}}{\sqrt{2}(a_2 - a_1)} e_2. \end{aligned}$$

This form allows us to read off the $n \rightarrow \infty$ limit directly except in the two special cases $\theta = \pm \frac{\pi}{2}$, as the second term on the right-hand side of (A.3) vanishes in the

limit unless $\sin \theta = -1$ (already excluded) or $\sin \theta = 1$:

$$(A.4) \quad \begin{aligned} \lim_{n \rightarrow \infty} \xi_n &= \lim_{n \rightarrow \infty} \xi_{m-1+n} \\ &= \frac{1}{\sqrt{2}(a_2 - a_1)} \begin{pmatrix} \sqrt{2}a_2\psi_{k-2}^{(m-1)} - \sqrt{2}a_1\psi_{k-1}^{(m-1)} \\ \sqrt{2}a_2\psi_{k-2}^{(m-1)} - \sqrt{2}a_1\psi_{k-1}^{(m-1)} \end{pmatrix}. \end{aligned}$$

Note that the two components of this limit vector are equal, so the “jump” in ψ disappears at the dyadic rational point x in question. This does not imply that ψ converges to a continuous function φ . While the values of $\psi^{(m-1+n)}(x)$ on the two intervals in the stage- $(m-1+n)$ partition at the immediate left of the x in question do move closer together as n increases until their difference becomes zero in the limit, values at higher-order dyadic rationals are not so well behaved: for example, we can see, by extending the treatment above to a 3×3 matrix $A' = \begin{pmatrix} \sqrt{2}a_3 & \sqrt{2}a_1 & 0 \\ 0 & \sqrt{2}a_2 & \sqrt{2}a_0 \\ 0 & \sqrt{2}a_3 & \sqrt{2}a_1 \end{pmatrix}$ iterating three intervals at the left of the point x instead of just two, that the value of $\psi^{(m-1+n)}(x)$ on the third interval added at the left of the original two generally grows without limit as $n \rightarrow \infty$ if the new eigenvalue $\sqrt{2}a_3$, corresponding to the eigenvector $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, is greater than 1, which is the case for $0 < \theta < \frac{\pi}{2}$.

As an example, the negative “peak” at $x = 1$ (see Figure 3(b)–(j)) can be obtained by using the starting vector $\xi_0 = \begin{pmatrix} \psi_{-1}^{(0)} \\ \psi_0^{(0)} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ (following the convention that $\psi^{(m)}(x)$ for $x < 0$ for all m , and thus $\psi_{-1}^{(0)} = 0$; see Figure 5): then the two components of $\lim_{n \rightarrow \infty} \xi_n$ are both equal to

$$\begin{aligned} -\frac{\sqrt{2}a_1}{\sqrt{2}(a_2 - a_1)} &= -\frac{1 - \cos \theta - \sin \theta}{2 \cos \theta} \\ &= \frac{1}{2} \left(1 + \tan \frac{\theta - \frac{\pi}{2}}{2} \right). \end{aligned}$$

Similarly, the peak at $x = 2$ in the same plots can be obtained by using the starting vector $\xi_0 = \begin{pmatrix} \psi_0^{(0)} \\ \psi_1^{(0)} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ (see Figure 5); it is

$$\begin{aligned} \frac{\sqrt{2}a_2}{\sqrt{2}(a_2 - a_1)} &= \frac{1 + \cos \theta - \sin \theta}{2 \cos \theta} \\ &= \frac{1}{2} \left(1 - \tan \frac{\theta - \frac{\pi}{2}}{2} \right). \end{aligned}$$

Note that these two peaks sum to 1 for all θ (excluding $\pm \frac{\pi}{2}$). The limiting value at $x = 3/2$ is also independent of θ . For this, use the starting vector $\xi_1 = \begin{pmatrix} \psi_1^{(1)} \\ \psi_2^{(1)} \end{pmatrix} = \begin{pmatrix} \sqrt{2}a_1 \\ \sqrt{2}a_2 \end{pmatrix}$ (see Figure 6(a)). This is just e_2 , so

$$\lim_{n \rightarrow \infty} \xi_n = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

The numerical values of $\lim_{n \rightarrow \infty} \psi^{(n)}(x)$, $x = 1, \frac{3}{2}, 2$ for the “movie reel” plots in Figure 3 are given below.

| Fig. 3 | θ | $x = 1$ | $x = \frac{3}{2}$ | $x = 2$ |
|--------|--------------------|---------|-------------------|---------|
| (b) | $-\frac{9\pi}{20}$ | -5.8531 | 0 | 6.8531 |
| (c) | $-\frac{2\pi}{5}$ | -2.6569 | 0 | 3.6569 |
| (d) | $-\frac{7\pi}{20}$ | -1.5826 | 0 | 2.5826 |
| (e) | $-\frac{3\pi}{10}$ | -1.0388 | 0 | 2.0388 |
| (f) | $-\frac{\pi}{4}$ | -0.7071 | 0 | 1.7071 |
| (g) | $-\frac{\pi}{5}$ | -0.4813 | 0 | 1.4813 |
| (h) | $-\frac{3\pi}{20}$ | -0.3159 | 0 | 1.3159 |
| (i) | $-\frac{\pi}{10}$ | -0.1882 | 0 | 1.1882 |
| (j) | $-\frac{\pi}{20}$ | -0.0854 | 0 | 1.0854 |
| (k) | 0 | 0.0000 | 0 | 1.0000 |
| (l) | $\frac{\pi}{20}$ | 0.0730 | 0 | 0.9270 |
| (m) | $\frac{\pi}{10}$ | 0.1367 | 0 | 0.8633 |
| (n) | $\frac{3\pi}{20}$ | 0.1936 | 0 | 0.8064 |
| (o) | $\frac{\pi}{5}$ | 0.2452 | 0 | 0.7548 |
| (p) | $\frac{\pi}{4}$ | 0.2929 | 0 | 0.7071 |
| (q) | $\frac{3\pi}{10}$ | 0.3375 | 0 | 0.6625 |
| (r) | $\frac{7\pi}{20}$ | 0.3800 | 0 | 0.6200 |
| (s) | $\frac{2\pi}{5}$ | 0.4208 | 0 | 0.5792 |
| (t) | $\frac{9\pi}{20}$ | 0.4606 | 0 | 0.5394 |

These values can be observed in the plots, though for θ near $\frac{\pi}{2}$ the cascade level plotted ($m = 8$) is not high enough to show a close approach to the limit.

Acknowledgements. We are grateful to Brian Treadway for insightful observations on the cascade implementations, and to Rune Kleveland for help with implementing the algorithm.

Most of the work in the present paper was done while O.B. visited Iowa in December of 1997, but it was interrupted at the end of the visit, when P.J. landed in the hospital after an accident.

REFERENCES

- [Akh65] N.I. Akhiezer, *The Classical Moment Problem and Some Related Questions in Analysis*, Oliver & Boyd, Edinburgh, 1965, translated by N. Kemmer from the Russian Классическая Проблема Моментов и Некоторые Вопросы Анализа, Связанные с Нею, Государственное Издательство Физико-Математической Литературы, Moscow, 1961.
- [BEJ99] O. Bratteli, D.E. Evans, and P.E.T. Jorgensen, *Compactly supported wavelets and representations of the Cuntz relations*, Applied and Computational Harmonic Analysis, to appear.
- [CoDa96] A. Cohen and I. Daubechies, *A new technique to estimate the regularity of refinable functions*, Rev. Mat. Iberoamericana **12** (1996), 527–591.
- [Coh92] A. Cohen, *Wavelets and digital signal processing*, Wavelets and their Applications, Jones and Bartlett, Boston, 1992, pp. 105–121.

- [CoHe92] D. Colella and C. Heil, *The characterization of continuous, four-coefficient scaling functions and wavelets*, IEEE Trans. Inform. Theory **38** (1992), 876–881.
- [CoHe94] D. Colella and C. Heil, *Dilation equations and the smoothness of compactly supported wavelets*, Wavelets: Mathematics and Applications (J.J. Benedetto and M.W. Frazier, eds.), Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1994, pp. 163–201.
- [CoRy95] A. Cohen and R.D. Ryan, *Wavelets and Multiscale Signal Processing*, Applied Mathematics and Mathematical Computation, vol. 11, Chapman & Hall, London, 1995.
- [DaLa92] I. Daubechies and J.C. Lagarias, *Two-scale difference equations, II: Local regularity, infinite products of matrices and fractals*, SIAM J. Math. Anal. **23** (1992), 1031–1079.
- [Dau92] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conf. Ser. in Appl. Math., vol. 61, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [Erd40] P. Erdős, *On the smoothness properties of a family of symmetric Bernoulli convolutions*, Amer. J. Math. **62** (1940), 180–186.
- [Hor95] L. Hörmander, *Lectures on harmonic analysis*, Dept. of Mathematics, Box 118, S-22100 Lund, 1995.
- [Hut81] J.E. Hutchinson, *Fractals and self similarity*, Indiana Univ. Math. J. **30** (1981), 713–747.
- [Pol89] D. Pollen, *Parametrization of compactly supported wavelets*, company report, AWARE, Inc., AD890503.1.4, 1989.
- [Pol90] D. Pollen, $SU_I(2, F[z, 1/z])$ for F a subfield of \mathbf{C} , J. Amer. Math. Soc. **3** (1990), 611–624.
- [Pol92] D. Pollen, *Daubechies’ scaling function on $[0, 3]$* , Wavelets: A Tutorial in Theory and Applications (C.K. Chui, ed.), Wavelet Anal. Appl., vol. 2, Academic Press, Boston, 1992, pp. 3–13.
- [Sol95] B.M. Solomyak, *On the random series $\sum \pm \lambda^n$ (an Erdős problem)*, Ann. of Math. (2) **142** (1995), 611–625.
- [Str96] G. Strang, *Eigenvalues of $(\downarrow 2)H$ and convergence of the cascade algorithm*, IEEE Transactions on Signal Processing **44** (1996), 233–238.
- [Vil94] L.F. Villemoes, *Wavelet analysis of refinement equations*, SIAM J. Math. Anal. **25** (1994), 1433–1460.
- [Wan95] Yang Wang, *Two-scale dilation equations and the cascade algorithm*, Random Comput. Dynam. **3** (1995), 289–307.
- [Wan96] Yang Wang, *Two-scale dilation equations and the mean spectral radius*, Random Comput. Dynam. **4** (1996), 49–72.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF OSLO, PB 1053 – BLINDERN, N-0316 OSLO, NORWAY

E-mail address: bratteli@math.uio.no

DEPARTMENT OF MATHEMATICS, THE UNIVERSITY OF IOWA, 14 MACLEAN HALL, IOWA CITY, IA 52242-1419, U.S.A.

E-mail address: jorgen@math.uiowa.edu